THE UNIVERSITY
of EDINBURGH

Applications of multi-way analysis for characterizing paediatric electroencephalogram (EEG) recordings

Eli W. Kinney-Lang

Doctor of Philosophy

The University of Edinburgh

October 2018

THE UNIVERSITY *of* EDINBURGH
School of Engineering

Dedication

For my family and Sara,

Summer without you is as cold as winter.

Winter without you is even colder.

Eli W. Kinney-Lang

October 2018

Abstract

This doctoral thesis outlines advances in multi-way analysis for characterizing electroencephalogram (EEG) recordings from a paediatric population, with the aim to describe new links between EEG data and changes in the brain. This entails establishing the validity of multi-way analysis as a framework for identifying developmental information at the individual and collective level. Multi-way analysis broadens matrix analysis to a multi-linear algebraic architecture to identify latent structural relationships in naturally occurring higher order (n-way) data, like EEG. We use the canonical polyadic decomposition (CPD) as a multi-way model to efficiently express the complex structures present in paediatric EEG recordings as unique combinations of low-rank matrices, offering new insights into child development. This multi-way CPD framework is explored for both typically developing (TD) children and children with potential developmental delays (DD), e.g. children who suffer from epilepsy or paediatric stroke.

Resting-state EEG (rEEG) data serves as an intuitive starting point in analyzing paediatric EEG via multi-way analysis. Here, the CPD model probes the underlying relationships between the spatial, spectral and subject modes of several rEEG datasets. We demonstrate the CPD can reveal distinct population-level features in rEEG that reflect unique developmental traits in varying child populations. These development-affiliated profiles are evaluated with respect to capturing structures well-established in childhood EEG. The identified features are also interrogated for their predictive abilities in anticipating new subjects' ages. Assessing simulations and real rEEG datasets of TD and DD children establishes the multi-way analysis framework as well suited for identifying developmental profiles from paediatric rEEG.

We extend the multi-way analysis scheme to more complex EEG scenarios common in EEG rehabilitation technology, like brain-computer interfaces. We explore the feasibility of multi-way modelling for interventions where developmental changes often pose as barriers. The multi-way CPD model is expanded to include four modes- task, spatial, spectral and subject data, with non-negativity and orthogonality constraints imposed. We analyze a visual

attention task that elucidates a steady-state visual evoked potential and present the advantages gained from the extended CPD model. Through direct multi-linear projection, we demonstrate that linear profiles of the CPD can be capitalized upon for rapid task classification sans individual subject classifier calibration.

Incorporating concepts from the multi-way analysis scheme with child development measured by psychometric tests, we propose the Joint EEG Development Inference (JEDI) model for inferring development from paediatric EEG. We utilize a common EEG task (button-press) to establish a 4-way CPD model of paediatric EEG data. Structured data fusion of the CPD model and cognitive scores from psychometric evaluations then permits joint decomposition of the two datasets to identify common features associated with each representation of development. Use of grid search optimization and a fully cross-validated design supports the JEDI model as another technique for rapidly discerning the developmental status of a child via EEG.

We then briefly turn our attention to associating child development as measured by psychometric tests to markers in the EEG using graph network properties. Using graph networks, we show how the functional connectivity can inform on potential developmental delays in very young epileptic children using routine, clinical rEEG measures. This establishes a potential tool complementary to the JEDI model for identifying and inferring links between the established psychometric evaluation of developing children and functional analysis of the EEG.

Multi-way analysis of paediatric EEG data offers a new approach for handling the developmental status and profiles of children. The CPD model offers flexibility in terms of identifying development-related features, and can be integrated into EEG tasks common in rehabilitation paradigms. We aim for the multi-way framework and associated techniques pursued in this thesis to be integrated and adopted as a useful tool clinicians can use for characterizing paediatric development.

Lay summary

Childhood reflects a critical developmental window for the brain, in which the brain is more receptive to medical intervention and recovery from sustained damage or injury. The complex and intricate brain functions associated with this improved recovery are a focus of great scientific interest, along with how to best introduce such interventions to the naturally diverse child population. The personal rates of cognitive and physical development in children, however, makes collective information on rehabilitation and intervention strategies less applicable to the general population. Therefore, there is a need for better techniques to capture, relay and interpret the distinct underlying developmental aspects of the brain present as children learn and grow.

Measuring brain activity under different conditions can be done readily and inexpensively using electroencephalography (EEG), which records differences in the timing, frequency and location of electrical activity produced by the brain via electrodes on the scalp. In this way, EEG recordings inherently include several dimensions of data, making EEG a naturally occurring 'higher order' dataset. This natural, multi-dimensional structure can be used to reveal important information about the brain and its response to tasks, changes and environments. To this end, this thesis focuses on exploiting the typical structures occurring within various EEG recordings to capture different aspects of child development.

In this thesis, the underlying relationships between child development and aspects of the EEG structure are explored under a variety of EEG recordings, including both at rest and during various stages of an engaging cognitive task. Complementary information gained from analyzing the brain from different perspectives, such as a network or as a higher order structure, is used to unearth key links in measuring a child's development. Measures like child age and cognitive test scores are then jointly considered in evaluating how well these techniques can capture development in a wide range of children, including in both typically developing children and those who may have suffered early developmental disruptions from disorders like epilepsy and paediatric stroke. We aim for the proposed techniques pursued in this thesis to be

integrated into the clinician's tool set for assessing child development.

List of Acronyms

- **BCI** - Brain-computer interface
- **CCP** - Contrast change paradigm
- **CHB-MIT** - Children's Hospital Boston-Massachusetts Institute of Technology (dataset)
- **CMI** - Child Mind Institute (dataset)
- **CORCONDIA**- Core consistency diagnostic
- **CP** - CPD/PARAFAC model
- **CPD** - Canonical polyadic decomposition (a.k.a. PARAFAC)
- **CST** - Cluster-span threshold
- **CWEOE** - Children with early onset epilepsy
- **ECoG** - Electrocorticography
- **EEG** - Electroencephalography
- **ICOH** - Imaginary part of coherency
- **JEDI** - Joint EEG Development Inference (model)
- **KKQ** - KidKINDL questionnaire
- **KNN** - K-nearest neighbor
- **MI** - Motor imagery
- **MMEC** - Muir Maxwell Epilepsy Centre (dataset)
- **MRCP** - Movement-related cortical potential
- **MST** - Minimum spanning tree

- **NFB** - Neurofeedback
- **PARAFAC** - Parallel factor analysis (a.k.a. CPD)
- **PLI** - Phase-lag index
- **PSI** - Phase-slope index
- **rEEG** - Resting-state EEG
- **SD** - Standard deviation
- **SDF** - Structured data fusion
- **SSVEP** - Steady-state visual evoked potential
- **SVM** - Support vector machine
- **WASI-II** - Wechsler Abbreviated Scale of Intelligence - Second Edition
- **WIAT-IIA** - Wechsler Individual Achievement Test - Second Edition, Abbreviated
- **WPLI** - Weighted phase-lag index

Notation

Notation

- $\mathcal{X}$ is a tensor.
- $\mathbf{X}$ is a matrix.
- $\mathbf{x}$ is a vector.
- $x_{ijk...n}$ are the elements of an N-way tensor $\mathcal{X}$
- $\mathbf{X}^{(n)}$ is the flattened, n-th way (i.e. mode or domain) of a tensor.
- $\circ$ is the outer product.
- $\otimes$ is the Kronecker product. Given matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{K \times L}$:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}B & \dots & a_{1j}B \\ \vdots & \ddots & \vdots \\ a_{i1}B & \dots & a_{ij}B \end{bmatrix} = \mathbf{T} \in \mathbb{R}^{IK \times JL} \tag{1}$$

- $\odot$ is the Khatri-Rao product (i.e. column-wise Kronecker product). Given matrices $\mathbf{C} \in \mathbb{R}^{I \times K}$ and $\mathbf{D} \in \mathbb{R}^{J \times K}$:

$$\mathbf{C} \odot \mathbf{D} = [c_1 \otimes d_1 \quad c_2 \otimes d_2 \quad \cdots \quad c_K \otimes d_K] = \mathbf{T} \in \mathbb{R}^{IJ \times K} \tag{2}$$

- $*$ is the Hadamard product (i.e. the elementwise multiplication). Given matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{I \times J}$:

$$\mathbf{A} * \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & \dots & a_{1j}b_{1j} \\ \vdots & \ddots & \vdots \\ a_{i1}b_{i1} & \dots & a_{ij}b_{ij} \end{bmatrix} = \mathbf{T} \in \mathbb{R}^{I \times J} \tag{3}$$

- $\mathbf{A}^T$ denotes the transpose of a matrix $\mathbf{A}$.
- $\mathbf{A}^\dagger$ denotes the Moore-Penrose pseduo-inverse of a matrix $\mathbf{A}$.
- *AD* is the average degree of a node.
- *BC* is the betweenness centrality of a node.
- *CC* is the clustering coefficient of a node.
- *D* is the diameter of a graph.
- *MD* is the maximum degree of a graph.
- *LF* is the leaf fraction of a graph.
- *V* is the degree variance of a graph.
- ρ is Pearson's correlation coefficient.
- τ is Kendall's tau coefficient

Acknowledgements

This thesis would not have been possible without the guidance and supervision of Dr. Javier Escudero, whose patience and diligent support of me was matched only by his relentless enthusiasm for my project. I am grateful to Dr. Nick Polydorides and Dr. Bonnie Auyeung for their conversations, guidance and wisdom as my second supervisors. I am thankful for the prudence and counsel of Dr. Richard Chin, without whom a significant portion of this thesis would not have been possible.

I would also like to thank Edinburgh Neuroscience, the Muir Maxwell Epilepsy Centre, and the Alberta Children's Hospital for supporting me and affording me unique opportunities and collaborations throughout this project. I would like to thank Dr. Michael Yoong and Dr. Matthew Hunter for their insights into the clinical workings of paediatric epilepsy, and their generous sharing of their available resources. I am thankful to Dr. Adam Kirton, Dr. Ephrem Zewdie of the University of Calgary for their consultation and enthusiasm about joint collaborations in paediatric BCI research, and in particular Abdullah Azeem for being willing to learn MATLAB from scratch to provide me with EEG spectral data.

A special mention goes to my colleagues and tensor-investigator-team of Ahmed Ebied and Dr. Loukianos Spyrou for many wonderful multi-dimensional conversations over the years, and to Dr. Keith Smith, Dr. Hamed Azami, and the other members of the Escudero Lab for their inspiration. Finally, I would like to thank my family, friends and better-half, Sara, for their kindness, support and encouragement throughout and beyond my doctoral research.

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own.

Eli W. Kinney-Lang

October 2018

Contents

Chapter 1

Introduction

1.1 Motivation

The work described in this thesis is motivated by a drive to better capture varying brain activity present in children as they develop, for practical incorporation into real-world, data-driven applications in medicine. The first step towards this goal was to examine state-of-the-art literature of a data-driven technology reliant on deciphering variable brain activity, e.g. brain-computer interfaces (BCI), and its accessibility for children. It became clear through the investigation that little research had explored how to adapt BCI technology to children, in part due to the complex, constantly evolving processes dominant in the developing brain. A large body of work therefore remains in terms of identifying, capturing and exploiting features in these changing paediatric brain processes, and is thus the majority focus for this thesis.

Better demarcation of the underlying properties of child development offer benefits ranging from assisting clinicians to opening up avenues for cutting-edge therapies and applications. Clinically, the time and effort required to appropriately identify, designate and track the developmental status of children and their response to early life insult or intervention is significant, with alternative strategies relatively limited due to heterogeneity in child development. Furthermore, the inherent variability of development has restricted access for children to fully embrace new rehabilitation strategies and applications, like BCI, which is rapidly advancing therapeutic opportunities for adults with impaired motor function. To this end, this thesis proposes characterizing electroencephalogram (EEG) recordings using advanced multi-way (tensor) analysis in conjunction with specific demographic information, like age, to identify and discern latent features depicting critical development-related

information. By capitalizing on the multi-modal data available in clinical EEG recordings, multi-way analysis is demonstrated on a variety of real-world datasets and applications. Evaluation across various paediatric populations, including both typical and abnormal developing children, contributes to the generalizability of the methods developed here. Successful progression from proof-of-concept to validation on real-world patient data will extend the characterization of paediatric EEG, and offer a quantitative framework which can support both clinicians and novel data-driven technologies.

1.2 Objectives

The main objective of the thesis is to establish multi-way analysis as a framework for characterizing paediatric development using EEG recordings. It is hypothesized that multi-way analysis will be able to capture and harness development-specific features from child EEG, and thus offer a means to adapt EEG-based applications and tasks for heterogeneous development in paediatric populations.

In order to achieve the main objective of the thesis, several partial objectives are identified. In particular, these partial objectives aim to:

- Establish the timeliness and relevance of the thesis work through a literature review.
- Demonstrate that multi-way analysis can identify development-related features using simple resting-state EEG data as a proof-of-principle.
- Extend the proof-of-principle to more complex EEG data useful in real-world conditions by applying the methods to steady-state visual evoked potentials (SSVEP).
- Utilize data fusion between a multi-way representation of a common motor task (button-press) relevant to BCI and other EEG based applications with standard psychological evaluations of child development in order to better characterize developmental states from child EEG.
- Exploit network analysis as a complementary technique to multi-way analysis for establishing pertinent information for characterizing child development.

1.3 Contributions

This thesis provides pioneering research on the development and modification of multi-way analysis for the purpose of interrogating paediatric EEG data for characterizing child development.

Evidence supporting the timeliness and importance of improving techniques for the characterization of paediatric EEG data is established in the state-of-the-art review on EEG-based BCIs for motor rehabilitation in children [1]. This review supports multi-way analysis as a potential solution to address specific barriers in current EEG analytic approaches, such as accounting for developmental variability across children [1].

The general applicability of the multi-way analysis framework is validated using a broad set of basic resting-state EEG data from a variety of typical and abnormal paediatric populations [2, 3]. These investigations establish the broad potential multi-way analysis has with respect to characterizing distinct paediatric populations.

The proposed framework is shown to be well suited under BCI application conditions using an SSVEP selection task, with the multi-way analysis adjusted in order to parse the additional task-related complexities of the SSVEP [4]. This work supports multi-way analysis in a BCI setting as a technique which does not require subject specific calibration of classifiers.

The proposed Joint EEG-Development Inference (JEDI) model then utilizes multi-way analysis to capture features in a common EEG task (button-press) and couples this with established psychometric scores (WASI-II, WIAT-IIA, and KKQ) via data fusion to infer child developmental status [5].

Finally, network analysis is illustrated as a complementary technique to multi-way analysis for identifying and characterizing markers of cognitive impairment present in child development [6]. The proposed network analysis is verified using a population of young children with epilepsy [6].

1.4 Structure of the Thesis

The remainder of the thesis is as follows:

Chapter 2: A summary of the background details and key elements of the thesis. This chapter discusses and provides context on the concepts underpinning the thesis. Details discussed include EEG and its recording, its changes during child development, multi-way analysis, the Canonical Polyadic Decomposition/Parallel Factor Analysis (CPD/PARAFAC) tensor model, and direct linear projection.

Chapter 3: A literature review of EEG-based BCI applications for motor rehabilitation in children. A detailed examination of real-world applications of BCI in children, which initially inspired the thesis research, is presented. The structured review supports the initial conception for the technical contributions presented in Chapters 4-6 based on unmet needs in BCI applications.

Chapter 4: Characterizing developmental feature profiles using multi-way analysis in paediatric resting-state EEG data. An initial proof-of-concept methodology is presented for utilizing multi-way analysis to capture development-related features in non-task based signals of paediatric rEEG. Resting-state data is interrogated as a 3-way (3-mode) tensor and evaluated on multiple paediatric datasets, including typical and abnormal developing children, in order to demonstrate the generalizability of the developed methods.

Chapter 5: Extensions of multi-way processing to 4-mode SSVEP data from childhood EEG. The multi-way framework from Chapter 4 is extended to task-based EEG signals, using an SSVEP two-choice selection paradigm. EEG data is selected to be as similar to a BCI SSVEP selection task as possible, and the multi-way framework is evaluated with respect to potential for incorporation into SSVEP-based BCI paradigms for children.

Chapter 6: Data fusion towards the Joint EEG Development Inference (JEDI) model. The JEDI model comprises the structured data fusion between an adapted multi-way framework representation of the common button-press EEG task and traditional psychometric scores of development. Details of the structured data fusion, the 4-mode button-press tensor and choice of psychometric scores are discussed in detail. Potential applications of the JEDI model are

covered, including its potential use in BCI for children.

Chapter 7: Network analysis of EEG as a complementary approach in characterizing impairment in children. A potential alternative methodology for identifying and characterizing markers of child development using graph network analysis is presented. In particular, cognitive impairment in children with early-onset epilepsy is investigated by analyzing networks of routine clinical child rEEG.

Chapter 8: Discussion, conclusions and future work. The main findings of the thesis are summarized, followed by discussion of limitations in the research. Potential avenues for future work is also presented, based on the novel developments produced in the thesis.

Chapter 2

Background

2.1 Introduction

With each subsequent generation of researchers, the boundaries between science and science fiction continue to bleed together, blending advancements across research disciplines. Less than three years after *The Matrix* [7] blew away the imagination of cinema goers with their depiction of a complete joining between the brain and a massive simulated world, a keystone paper describing real-world brain-computer interface (BCI) designs was published [8]. In this foundational work, Wolpaw *et al.* helped to establish the modern field of BCI research, where concepts from electrical engineering, computer science, physiology, psychology and neuroscience equally contributed to new scientific developments.

One core pillar to this new technology is the ability to identify, decipher and encode the information contained in the electric signals produced by the brain [8]. This pillar depends on a multi-disciplinary understanding of several topics, like the acquisition and processing of brain signals and their relationship to underlying neurophysiological phenomena [8]. This chapter aims to summarize the key concepts which are fundamental to the thesis and provide a frame of reference for its contributions to the field of BCI research.

2.2 Electric Life Orchestra: Neurons and Electrical Activity in the Brain

The brain is composed of billions of cells called neurons. A neuronal cell is comprised of a nucleus inside a cell body, referred to as the soma [9]. The soma is surrounded by

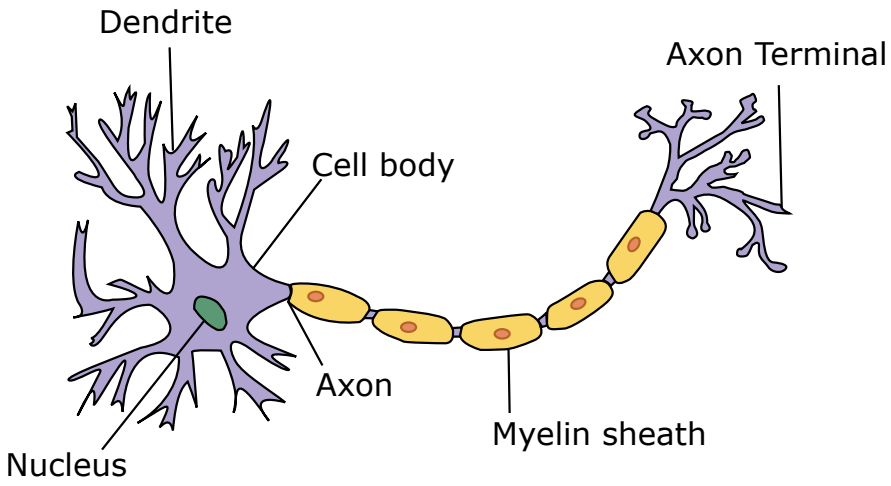

Figure 2.1: Illustration of a neuron and its various parts, labelled. This picture was adapted from a publicly available, creative-commons license free repository (1).

information-receiving filopodia, called dendrites, in addition to a single long axon wrapped in myelin sheath [9]. The dendrites and axon function as a highway for neurons to pass electrochemical impulses, and by extension information, from one neuron to another [9]. Electrochemical signals are taken in by the dendrites and passed on through the far end of the neuron at the axon terminal synapse [9]. The synapse functions as the port for inter-neuronal interfacing between the axon and dendrites of separate neurons. Figure 2.1 provides an illustration of the basic neuron.

Neuronal activity is rhythmic and repetitive, producing oscillatory brain activity [9, 10]. The oscillatory behaviour is driven by mechanics at the individual neuron level as well as through group interactions between neurons. In individual neurons, the rhythmic pattern is generated by action potentials, e.g. neuron firing. The action potential is generated by rapid depolarization along the axon, followed by re-polarization of the membrane potential [9]. This rapid depolarization/re-polarization rhythmically cascades along the axon, propagating the electrochemical signal from near the cell body to the synapse at the axon terminal [9]. At the group level, sets of neurons can simultaneously fire together, resulting in a synchronized pattern of neuron activation [9].

The rhythmic electrical activity in neurons are the driving force behind the brain. The synchronized, active transmission of electric impulses generated by millions of neurons are responsible for the brain's activities. This synchronized neuronal excitation produces the

impulses behind thoughts, movement, emotions and sensory input [9–11]. As such, capturing and understanding this orchestra of electric activity in the brain provides a window into comprehending the brain's complex biological processes.

2.3 The Electroencephalography: EEG

Early in the 20th century Hans Berger, a German psychiatrist, pioneered the recording of electrical activity in the human brain using a technique he coined 'the electroencephalogram' (EEG) [10]. EEG offered a new tool to clinicians, physiologists and scientists alike to non-invasively record and monitor neuronal activation in the human brain, via electric potential discharges. The electrical discharges of individual neurons were too small to detect effectively. However, through EEG, the synchronous activity of millions of neurons firing together could be recorded using simple scalp electrodes. Thus began the study of links between synchronous neuronal activity measured by EEG and the corresponding thoughts, behaviours, actions and diseases of patients being recorded.

2.3.1 EEG Brain Rhythms

Rhythmic waveforms in the EEG confer developmental and pathological significance. EEG recordings reflect insight into different states of the brain, e.g. alert wakefulness, drowsiness, etc., as well as brain processes such as attention, emotion, etc. [10]. Research into EEG has identified specific rhythmic frequencies which correspond to such phenomena. These waveforms fall into the several frequency bands, named the delta, theta, alpha, mu, beta, and gamma bands. A brief introduction to these classical frequency bands, and their association with specific brain processes follows.

The lowest frequency band is the delta band, covering the 1-3 Hz range. The delta band is associated with drowsiness, and becomes prominent as a person progress through the stages of sleep. While delta waves can be detected during normal wakefulness, a dominant delta frequency present in a vigilant adult can indicate possible focal, regional or generalized cerebral dysfunction. Due to the low frequency range of delta, brain processes of interest in this range may be difficult to distinguish from other electrophysiologic activity.

Similar to the delta band is the theta band, prominent in the 4-7 Hz range. Theta waves also notably play a prevalent role during drowsiness and sleep. Theta waves are typically absent, however, during the deepest stages of sleep. In very young children, cortical theta activity has been observed as the dominant natural rhythm [12]. This functionality of the theta band disappears as a child develops, and is absent in older children and adults.

The alpha band was the first neural oscillation range studied by Berger through EEG. Alpha waves are prominent from 8-12 Hz, and can be observed when the eyes are closed. The alpha rhythm is detected in the occipital region of the brain (rear of the head), and may be referred to as the posterior dominant rhythm. Variants of the alpha rhythm have been recorded, such as frontal alpha activity and temporal alpha, which are characterized by synchronous activity in the brain in the alpha frequency range, over those specific brain areas. A significant amount of research has evaluated how the alpha frequency behaves under different physiological conditions and brain activity due to its prominence across individuals.

Activity in the alpha band over the central motor cortices of the brain is referred to as the mu band. Mu waves can be detected unilaterally or bilaterally over the central brain regions. Activity in the mu band fluctuates with movement, imagined movement, and somatosensory stimulus. During such events, the mu frequency range attenuates its signals in the motor cortex region. This attenuation is spatially distinct from the alpha rhythm over the occipital brain. Thus, the mu rhythm is useful in discerning movement related phenomena in EEG recordings [10].

The beta band is designated for frequencies between 13-30 Hz. Beta waves reflect a wide variety of physiological processes, and are associated with the development of the frontal cortex and active cognition [11]. Beta waves, like mu, can also be detected over the central regions of the brain corresponding to the motor cortex [10, 11, 13]. The beta rhythm in this region is heavily implicated during motor activities, and is modulated during both the planning and execution phase of grasping movements [13].

Frequencies above 30 Hz belong to the gamma band [10]. This region is typically split into low and high gamma at 30-70 Hz and > 70 Hz, respectively. Historically the brain processes and information from the gamma band has been challenging to interpret and analyze using

scalp EEG. Consequently, this has led to less research and exploration of this region. However, advances in signal analysis and processing has given a surge to research aiming to decode this frequency band for EEG analysis [14, 15].

2.3.2 EEG Hardware and the 10-20 Montage

A typical EEG recording system requires a cap to hold the electrodes. Electrodes are systematically spaced over the scalp in a standardized system, such as the 10-20 system [10, 11, 16]. The 10-20 system is a recognized method to describe the location and identify the position of scalp electrodes for EEG studies and research [10, 11]. The "10-20" nomenclature represents a standard for the intervals of measurement (either 10% or 20%) when positioning electrodes over the anterior-to-posterior dimensions of the brain. Electrode placement ranges from the bridge of the nose (naison) to the prominent bump on the back of the head at the lowest point of the skull (inion) and between the two ear positions (auricular) [10]. The 10-20 system then assigns each electrode site a specific letter and number. The letter represents which lobe of the brain the electrode site overlies, while the number indicates where the electrode is with respect to the left (odd number) or right (even number) brain [10]. For example, the electrode O2 refers to the electrode over the right occipital lobe of the brain.

The 10-20 system is not the only available montage for EEG electrodes. Often, more dense electrode placements have been devised, subdividing the 10-20 system further to include additional electrode placements. In the typical 10-20 system, approximately 20 electrodes are used to cover the brain. More dense electrode montages may use the base 10-20 system, but with greater numbers of electrodes placed between the standard 10-20 electrode sites. Figure 2.2 illustrates such a system, where the 10-20 montage was used as a basis for a dense EEG network of 64 electrodes. An example of an actual EEG electrode cap on a person is also shown in Figure 2.3.

Using the principle of differential amplification, pairs of electrodes are used to record voltage differences between an active scalp electrode site and a reference electrode, either nearby or distant. From the recorded differences, the aggregated potential produced by electrical activity in the brain is measured. The potential recorded from each of these electrodes is then displayed

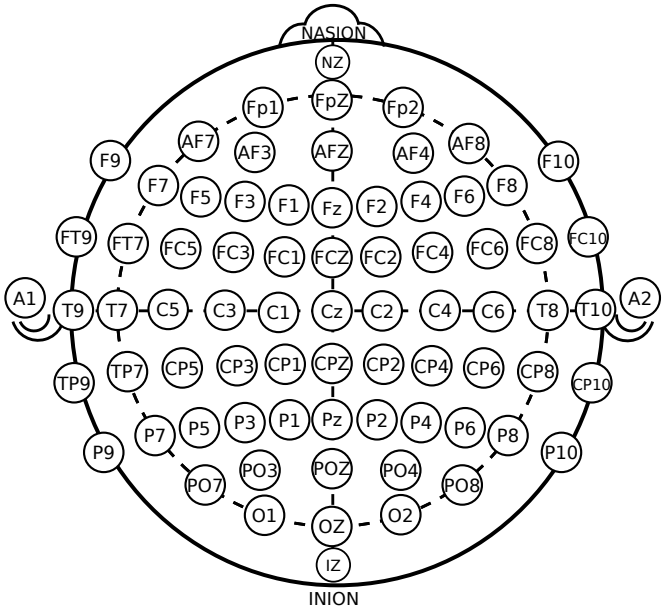

*Figure 2.2: An illustration of the 10-20 system for EEG placement, adapted to accommodate the dense 64-electrode EEG montage using a topographic representation of the scalp. Electrode letters correspond to different brain regions: Pre-frontal (**Fp**); Frontal (**F**); Temporal (**T**); Central (**C**); Parietal (**P**); Occipital (**O**). Additional channels between these major lobes are also marked, e.g. **AF** represents channels between **Fp** and **F**. The image provided is part of the public domain and uploaded from an open repository (2).*

Figure 2.3: Example of a (handsome) person wearing a standard EEG cap.

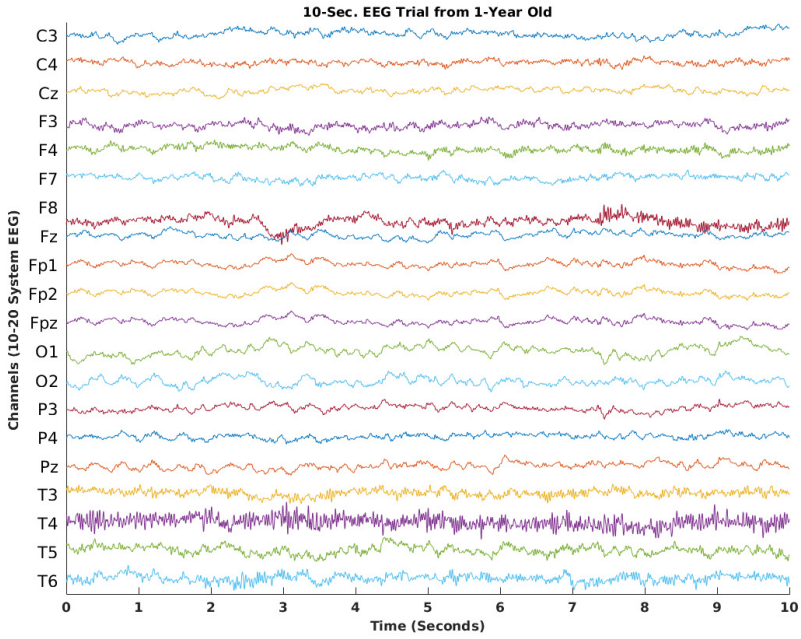

Figure 2.4: An example of a standard EEG time-series. Different streams of data on the y-axis represent the distinct electrode channels used in the EEG, which cover the scalp in the standard 10-20 system. The x-axis shows the length of time of recording data from each channel. Signal amplitude in each time-series recording is on the scale of microvolts (μV).

on the vertical axis of an EEG graph, with time represented on the horizontal axis. This gives a real-time display of ongoing cerebral activity in the brain as seen in Figure 2.4. Due to this fine temporal resolution, a predominant utility arising from EEG is the evaluation of dynamic cerebral functioning [10,11]. EEG offers insight to the rapid changes in physiological processes at the millisecond level, and thus has proved invaluable as a tool to study both typical brain rhythms and deviations from normal patterns associated with diseases [10, 11, 17].

EEG has some relevant physical limitations in assessing brain processes. The specific brain activity of interest can by masked by other electrical activity picked up by the EEG recording. For example, electric activity generated by the body, brain processes not of interest or even the environment can produce strong enough currents to obfuscate the signals of interest [10]. Additionally, EEG potentials generated from the source of interest in the brain must pass through multiple layers of biological tissue (e.g. brain, cerebral-spinal fluid, the meninges, the skull and the skin) to be recorded by the surface scalp electrodes. As such, when the signal potential of interest propagates through these various tissues, the tissues function as

biological filters which reduce the signal amplitude disperse the signal, potentially causing it to be picked up by electrodes distant from the original source. Other biological processes, e.g. scalp muscle movement, eye blinks, etc., can also generate significant potentials which have greater amplitude than the cerebral activity of interest, causing the brain process to be indistinguishable. Fortunately, research has identified many of the distinguishing characteristics of such artefacts, leading them to be identified by both human observers, and automated processes [10, 18–20] for improved EEG analysis.

This thesis aims to exploit EEG signals as a discrete measure which reveals the developmental state of a child's brain [21,22], with the goal to establish a new framework to better characterize the changing information present in paediatric EEG recordings.

2.3.3 Verifying Resting-state EEG

Recording EEG in a 'resting-state' is a common concept used throughout EEG literature [23–26]. Although this implies the brain is at rest, this is a misnomer. Resting-state EEG (rEEG) does not only include recordings from when patients are asleep (or otherwise incapacitated). Instead, resting state activity commonly refers to measuring the brain when no (clear) selective information processing occurs [27]. Typically this is when a patient is instructed to sit or lay quietly, with no outside stimulus presented, such as instructions for a specific task. While discrete electric processes are still happening in the brain during this recording window, the lack of specific required processing is taken to be the resting or 'default' mode of the brain [24]. The default state provides insight into brain oscillations occurring during this non-specific processing time for an individual. The above conditions have been widely accepted in the scientific community as an adequate representation of resting-state brain activity.

To mitigate potential effects from different non-specific brain processes occurring in rEEG recordings, this thesis uses two approaches. First, multiple rEEG recordings are averaged together for each individual, to provide a general picture of their specific resting-state processes, and mitigate potential outlying brain activities (e.g. specific motor imagery thoughts/changes in attention etc.). Additionally, where possible, rEEG trials were collected directly following initially placing on the EEG cap. This provided a similar activity condition across the patients, and guaranteed that the patients were awake during the resting-state

recordings. In cases where longer rEEG recordings occurred, e.g. data from the Children's Hospital Boston-MIT in Section 4.2.1, a large number of rEEG trials were averaged together to render a holistic representation of the resting-state processes for each subject.

2.3.4 Child Development and the EEG

From birth to adulthood the reorganization and development of new connections in the brain leads to an evolving set of neurophysiological properties [1, 22, 28]. Consequently, EEG rhythms corresponding to the brain's activity also vary throughout a child's development [12, 21, 28]. Recorded paediatric EEG signals are significantly altered compared to the EEG rhythms seen in adult recordings, indicating the need to separately analyze paediatric EEG. For example, the well studied EEG alpha rhythm changes throughout child development [12]. Migration of alpha rhythm progresses from approximately 6-9 Hz during childhood to 8-12 Hz in adulthood [12, 21, 28]. Shifts in EEG frequency bands is a well-established phenomena corresponding to the maturation and developmental progression of the brain throughout childhood [12, 21, 22, 28, 29].

Shifting properties of the classic frequency bands of the EEG happen alongside shifts in the spatial distribution of the EEG. The accompanied spatial changes reflect the physical reorganization of the brain during child development [28, 29]. Again using the alpha rhythm as an example, the associated regional connectivity in the occipital area of the brain is significantly altered between children and adults [29]. This again highlights how development affects brain signals recorded by the EEG at various levels. These shifts in EEG signal recordings from young children are further confounded by strong brain processes which are not of interest (commonly referred to as 'high background noise') in the child's EEG [12, 21, 22]. The influence of background physiological processes reflect how dominating low frequency bands, such as theta, are prevalent in young children and how they may affect EEG recordings in that population. Further the young brain has reduced power in higher frequency bands typically dominant in mature brains, like beta. The prominence of low electrophysiological processes affecting EEG recordings persists throughout childhood until adolescence, resulting in potentially obfuscated signals of interest [12].

These examples illustrate some of the challenges which face gaining useful information from

EEG recordings in children. However, despite these barriers, EEG should not be disregarded, as its affordability, accessibility and non-invasive nature still position it as a convenient tool for readily inferring a wide array of information about brain activity and development [21, 30–32]. This thesis proposes to address these barriers by utilizing multi-way analysis for inferring and discerning relevant information from paediatric EEG data.

2.4 Matrix Analysis and Factorization

As a preface to multi-way analysis, a brief introduction to matrix (2-way) analysis and factorization is presented here.

Matrix factorization is a critically important and unifying topic in signal processing and linear algebra, which has been fundamental to a significant spectrum of applications [33]. In general terms, matrix factorization takes a given set of data which can be represented by a matrix $\mathbf{X}$, such as a system of linear equations or a group of observations, and decomposes the data via linear algebra manipulations into a set of matrices, e.g.:

$$\mathbf{X} = \mathbf{AB} \tag{2.1}$$

where $\mathbf{A}$, and $\mathbf{B}$ are some result matrices from the particular decomposition of $\mathbf{X}$. These manipulations are chosen in a specific way to complement a particular goal, such as reducing, modelling or interpreting the information from the data matrix. Some common matrix factorization methods include the principal component analysis (PCA) and non-negative matrix factorization (NNMF) [33–35].

PCA is one of the oldest and most widely used techniques for reducing the dimensionality of a large dataset in order to interpret key aspects underlying the data matrix [34, 35]. The goal of PCA is to reduce the dimensionality of the data while preserving as much variability as possible in so called principal component variables [34]. Retaining the statistical information amounts to finding new, uncorrelated variables which are linear functions of those present in the original dataset, and which successively maximize variance [34]. These new variables, e.g. the principal components, can be considered as effectively solving a specific

eigenvector/eigenvalue problem [34]. The standard example for PCA is as follows.

Let $\mathbf{X}$ be a data matrix, with p numerical observations of n entities, that is $\mathbf{X} \in \mathbb{R}^{n \times p}$. The PCA aims to find a linear combination of the columns of $\mathbf{X}$, e.g. $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$, with maximum variance [34]. These linear combinations are given by:

$$\sum_{j=1}^{p} a_j \mathbf{x}_j = \mathbf{Xa} \tag{2.2}$$

where $\mathbf{a}$ is a vector of constants [34]. The variance for any linear combination in 2.2 is given by:

$$var(\mathbf{Xa}) = \mathbf{a}^T \mathbf{S} \mathbf{a} = \lambda \mathbf{a}^T \mathbf{a} = \lambda \tag{2.3}$$

for a sample covariance matrix $\mathbf{S}$, and the restriction that $\mathbf{a}$ must be a unit-norm eigenvector, e.g. $\mathbf{a}^T \mathbf{a} = 1$, with corresponding eigenvalue λ [34]. By enforcing the unit-norm restriction on $\mathbf{a}$, an orthonormal set of vectors can be found for the eigenvectors of $\mathbf{S}$ which function to successively maximize variance subject to being uncorrelated with previous linear combinations, and thereby solve our initial goal in PCA [34].

Alternatively, the PCA can be understood as a singular value decomposition of the column-centred data matrix $\mathbf{X}$. For k principal components, this is represented as:

$$\mathbf{X} = \mathbf{U} \Sigma \mathbf{V}^T \tag{2.4}$$

where $\mathbf{U} \in \mathbb{R}^{n \times k}$, $\mathbf{V}^T \in \mathbb{R}^{k \times p}$ are column-wise orthonormal and $\Sigma \in \mathbb{R}^{k \times k}$ is a diagonal matrix of k non-negative singular values. Each principal component on the diagonal of Σ corresponds to a specific column and row of $\mathbf{U}$ and $\mathbf{V}^T$ respectively. Figure 2.5 illustrates the PCA with respect to the SVD representation.

Critically, the PCA relies on orthogonality to find its unique solution. Without this constraint the results would be susceptible to translational variants between the principal components and subsequently would be satisfied by any number of solutions [34]. Therefore to retain the interpretable meaning and correspondence between the principal components and the associated loading vectors in $\mathbf{U}$ and $\mathbf{V}$, PCA requires strict use of orthogonality in the analysis.

NNMF is a technique which is similar to the PCA in its goal. The NNMF aims to decompose a given matrix into (usually) two constituent matrices, with the property that all elements are strictly non-negative across the collection of matrices [35, 36]. By ensuring non-negativity for all elements, the NNMF permits easier interpretation of the data being processed, especially when negative elements are unnatural, e.g. in processing some biological signals [35]. As an example, again let $\mathbf{X} \in \mathbb{R}^{n \times p}$. Then the NNMF of $\mathbf{X}$ can be written as the product of matrices $\mathbf{W}$, $\mathbf{H}$:

$$\mathbf{X} = \mathbf{WH} \tag{2.5}$$

where all elements of $\mathbf{W}$ and $\mathbf{H}$ are non-negative. The matrix multiplication can be understood as finding the column vectors of $\mathbf{X}$ as linear combinations of column vectors in $\mathbf{W}$ as multiplied by the columns of $\mathbf{H}$. In other words, each column of $\mathbf{X}$ is calculated as:

$$\mathbf{x}_i = \mathbf{W}\mathbf{h}_i \tag{2.6}$$

where $\mathbf{x}_i, \mathbf{h}_i$ are the i-th columns of $\mathbf{X}$ and $\mathbf{H}$, respectively. Notice, however, that the given factorization is not unique [35, 36]. A matrix and its inverse could be used to transform the factorization matrices of $\mathbf{W}$, and $\mathbf{H}$ such that they remain non-negative, but correspond to a

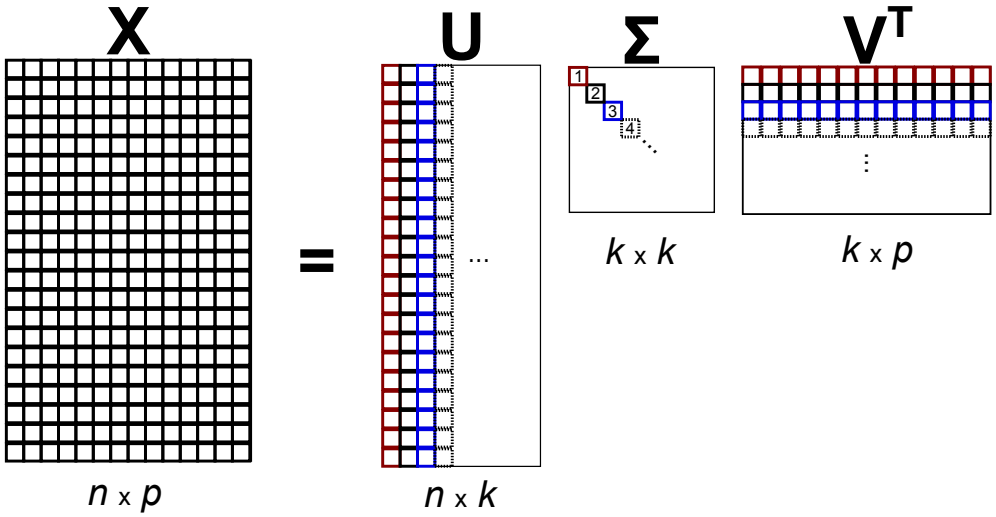

Figure 2.5: Visual representation of the PCA as outlined in 2.2. The first four principal components in Σ are given on the diagonal, and visually coded to correspond to the associated values in $\boldsymbol{U}$ and $\boldsymbol{V}^T$.

different scale or permutation, e.g. $\hat{\mathbf{W}}, \hat{\mathbf{H}}$. Written out, that is:

$$\mathbf{WH} = \mathbf{WBB}^{-1}\mathbf{H} \tag{2.7}$$

for some invertible matrix $\mathbf{B}$, with $\hat{\mathbf{W}} = \mathbf{WB}$, and $\hat{\mathbf{H}} = \mathbf{B}^{-1}\mathbf{H}$. Therefore, additional constraints like sparsity or orthogonality need to be introduced to the NNMF to guarantee its uniqueness [35, 36].

Matrix factorization like PCA and NNMF offers simple but effective methods for analysis of 2-way arrays and data matrices. Higher-order arrays (e.g. $>$ 2-ways), however, can be more problematic for matrix factorization, with information potentially lost as a result of the 2-dimensional matrix representation [33, 37]. Thus, a different set of tools needs to be considered when analyzing these higher-order data arrays.

2.5 Multi-way Analysis

Analogous to how linear algebra forms the basis for matrix analysis techniques, multi-linear algebra is the cornerstone on which multi-way analysis is built. Multi-linear algebra explores how to naturally extend the standard matrix analysis techniques above to higher-order data and presents the framework using tensor models and decompositions [38–41].

2.5.1 Brief Introduction to Notation

Multi-way data is typically comprised of three or more ways, also called modes or domains. Tensors retain the informative structural associations between these higher-order modes often lost using two-way matrix models [40, 41]. A tensor is designated by a calligraphic upper-case letter (e.g. $\mathcal{X} \in \mathbb{R}^{I \times J \times \cdots \times N}$). Matrices can be considered as 2-way tensors, and are represented by a bold upper-case letter (e.g. $\mathbf{X} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_J] \in \mathbb{R}^{I \times J}$). Vectors are 1-way tensors designated by a bold lower-case letter (e.g. $\mathbf{a}_j$). The n-th way flattened matrix of a tensor $\mathcal{X}$ is written as $\mathbf{A}^{(n)}$ and is the corresponding mode-n slice of the whole tensor. For example, given a three-dimensional tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ with elements x_{ijk}, the mode-1 frontal slice of $\mathcal{X}$ is given by $\mathbf{X}^{(1)} \in \mathbb{R}^{I \times JK}$, the mode-2 lateral slice by $\mathbf{X}^{(2)} \in \mathbb{R}^{J \times IK}$ and mode-3 horizontal slice by $\mathbf{X}^{(3)} \in \mathbb{R}^{K \times IJ}$. Figure 2.6 gives a side-by-side representation of vector, matrix and

tensor objects.

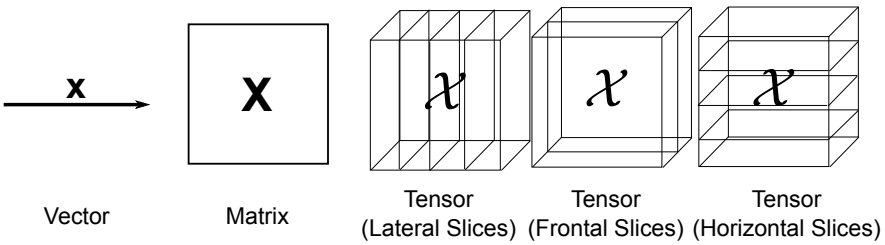

Figure 2.6: Graphic representation of vector, matrix and tensor objects. The tensor objects are shown with corresponding lateral, frontal and horizontal matrix slices.

The Kronecker product of two given matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{K \times L}$ is denoted by $\mathbf{A} \otimes \mathbf{B}$ [42]. The result is a matrix $\mathbf{T} \in \mathbb{R}^{IK \times JL}$ with:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}B & \dots & a_{1j}B \\ \vdots & \ddots & \vdots \\ a_{i1}B & \dots & a_{ij}B \end{bmatrix} = \mathbf{T} \in \mathbb{R}^{IK \times JL} \tag{2.8}$$

which is equivalent to

$$\mathbf{T} = [a_1 \otimes b_1 \quad a_1 \otimes b_2 \quad a_1 \otimes b_3 \cdots a_J \otimes b_{L-1} \quad a_J \otimes b_L] \tag{2.9}$$

The Khatri-Rao product is considered as the column-wise Kronecker product. Given two matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ and $\mathbf{B} \in \mathbb{R}^{J \times K}$, their Khatri-Rao product is given as $\mathbf{A} \odot \mathbf{B}$ [42]. This is equivalent to:

$$\mathbf{A} \odot \mathbf{B} = [a_1 \otimes b_1 \quad a_2 \otimes b_2 \quad \cdots \quad a_K \otimes b_K] = \mathbf{T} \in \mathbb{R}^{IJ \times K}$$

where $\mathbf{T}$ is the result matrix of interest from the Khatri-Rao [42].

The Hadamard product is the elementwise matrix product between two matrices of similar size, e.g. $\mathbf{A} * \mathbf{B}$ where $\mathbf{A}$ and $\mathbf{B}$ both are of size $\mathbb{R}^{I \times J}$ [42]. Then the Hadamard product is defined as:

$$\mathbf{A} * \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & \dots & a_{1j}b_{1j} \\ \vdots & \ddots & \vdots \\ a_{i1}b_{i1} & \dots & a_{ij}b_{ij} \end{bmatrix} = \mathbf{T} \in \mathbb{R}^{I \times J} \tag{2.10}$$

The Kronecker, Khatri-Rao and Hadamard products have several useful properties [42] including:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \tag{2.11}$$

$$(\mathbf{A} \otimes \mathbf{B})^{\dagger} = \mathbf{A}^{\dagger} \otimes \mathbf{B}^{\dagger} \tag{2.12}$$

$$(\mathbf{A} \odot \mathbf{B})^{T}(\mathbf{A} \odot \mathbf{B}) = \mathbf{A}^{T}\mathbf{A} * \mathbf{B}^{T}\mathbf{B} \tag{2.13}$$

$$(\mathbf{A} \odot \mathbf{B})^{\dagger} = ((\mathbf{A}^{T}\mathbf{A}) * (\mathbf{B}^{T}\mathbf{B}))^{\dagger}(\mathbf{A} \odot \mathbf{B})^{T} \tag{2.14}$$

where $\mathbf{A}^{\dagger}$ denotes the Moore-Penrose pseudoinverse of $\mathbf{A}$ [42].

An *N*-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times \cdots \times N}$ is considered rank-one if it can be written as the outer product of *N* vectors [42], i.e.,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}$$

where the symbol '$\circ$' denotes the vector outer product. The *M* groupings in Figure 2.7 illustrates a third-order rank-one tensor, e.g. *M(1)* corresponds to $\mathbf{a}_1 \circ \mathbf{b}_1 \circ \mathbf{c}_1$ for the tensor $\mathcal{X}$.

2.5.2 Links to EEG

EEG data readily supports multi-way models and analysis due to its inherent higher-order structure. In particular, implicit relationships exist between the EEG time-series ($[Temporal]$ data), the EEG electrode channels ($[Spatial]$ data) and the recorded EEG frequencies ($[Spectral]$ data) [43], providing a natural 3rd order data array. To clarify with an example, let a given EEG consist of the aforementioned distinct $[Temporal]$, $[Spatial]$, and $[Spectral]$ data modes. Then multi-way tensor analysis informs on the inherent relationships between these modes without destroying their underlying structural relationships [41]. Traditional two-way matrix analysis on the other hand would inform on some two-set combination of these domains, such as $[Spatial]$-$[Temporal]$ or $[Spatial]$-$[Spectral]$ combinations, losing information associated in the mode not included in the pairwise set [41]. Thus to more completely capture the complex structural information and relationships useful in EEG, multi-way analysis is preferable [41].

Advantages of tensor analysis have been explored in a variety of adult EEG data, with benefits demonstrated for EEG-dependent technologies such as BCI (see Chapter 3 for an in depth review) [38, 41, 44]. By exploiting the multi-way connections inherent in paediatric EEG, we hypothesize that tensor analysis offers a framework to capture underlying developmental links present throughout childhood.

2.5.3 CPD/PARAFAC Tensor Model and Decomposition

Canonical Polyadic Decomposition (CPD) [45]), also known as Parallel Factor Analysis (PARAFAC; used interchangeably throughout this thesis) [46], is a common tensor model used to represent higher-order data [38, 40, 42]. The modelled data is subsequently decomposed into its primary components, as shown in Figure 2.7. CPD/PARAFAC decompositions can thus be considered as the multi-linear algebra generalization of PCA for multi-way arrays (e.g. tensors) [42]. The CPD/PARAFAC analysis provides a model of the structural relationships between the different ways (e.g. modes) present in a tensor, providing insight to potential latent relationships [42, 47].

The CPD/PARAFAC (CP) model decomposes a multi-way tensor $\mathcal{X}$ into the sum of component rank-1 tensors coupled with a super-diagonal core [42, 45, 46]. Here we present a general *N*-mode example of the CP model, alongside a pictorial representation of the CP decomposition of a 3-way tensor in Figure 2.7.

Let $\mathcal{X} \in \mathbb{R}^{I \times J \times K \times \cdots \times N}$ be an *N*-way tensor. Then a general *R*-component CP model decomposition is defined as:

$$\mathcal{X} \approx \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \cdots \circ \mathbf{n}_r \tag{2.15}$$

for $r = 1, 2, .., R$ with $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$,$\mathbf{c}_r \in \mathbb{R}^K$, ... , $\mathbf{n}_r \in \mathbb{R}^N$ [42].

The CP model can alternatively be written as:

$$x_{ijk...n} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} \ldots m_{nr} + e_{ijk...n} \tag{2.16}$$

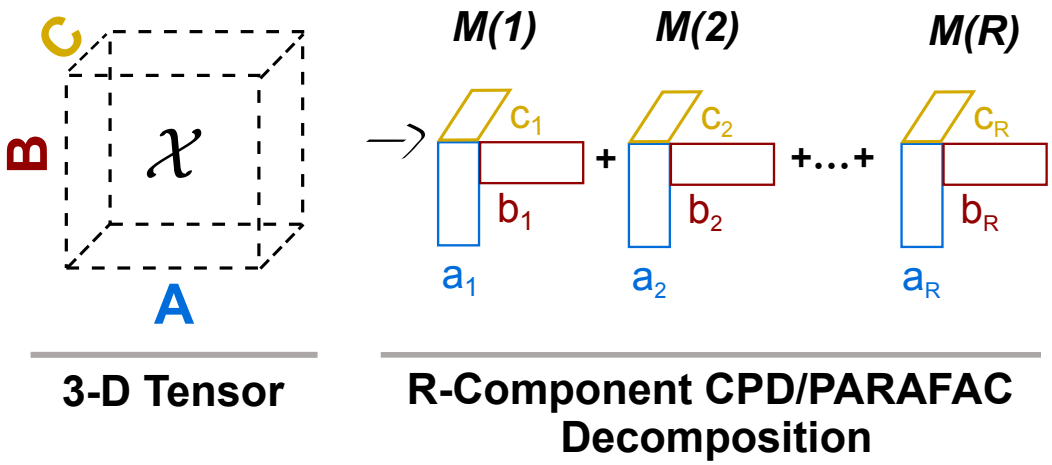

*Figure 2.7: Graphical representation of the CP decomposition for a generic 3-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ into R components. Relationships corresponding between the different modes **A**, **B**, **C** occur strictly between components with the same subscripts as shown in the M groupings, i.e. a_1 with b_1 and c_1, a_2 with b_2 and c_2, etc.*

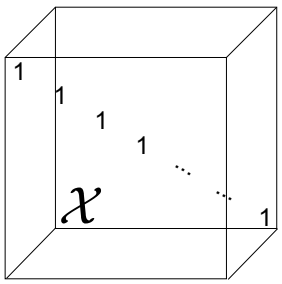

Figure 2.8: Example of a super-diagonal tensor. Values are only present on the diagonal of the tensor, and zero elsewhere.

for $i = 1, ..., I; \quad j = 1, ..., J; \quad k = 1, ..., K; \quad \ldots n = 1, ..., N$ and $r = 1, ..., R$ with $x_{ijkn}, a_{ir}, b_{jr}, c_{kr}, \ldots, m_{nr}$ and $e_{ijk\ldots n}$ as elements of $\mathcal{X}$, domains $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$, $\mathbf{C} \in \mathbb{R}^{K \times R}, \ldots, \mathbf{M} \in \mathbb{R}^{N \times R}$ and residual $\mathcal{E} \in \mathbb{R}^{I \times J \times K \times \cdots \times N}$ [42].

In both representations (2.15) and (2.16), a super-diagonal core tensor is used for the model, with the core visualized in Figure 2.8 [42].

Uniqueness of the CP model can be guaranteed under mild conditions, thereby retaining meaning in the low-rank factor matrices obtained from the CP decomposition [38, 42]. For example, let $\mathcal{X} \in \mathbb{R}^{I \times J \times \cdots \times N}$ be an N-way tensor with mode-n domain matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots, \mathbf{N}$. Also, let k_n be the rank for each domain matrix $\mathbf{A}, \mathbf{B}$, etc. for $n = 1, \ldots, N$. Then the decomposition of $\mathcal{X}$ into R-components is guaranteed to be unique under the

sufficient condition [42]:

$$\sum_{n=1}^{N} k_n \geq 2R + (N-1) \tag{2.17}$$

2.5.4 Tucker Model

The CP model can be considered as a special case of a more general tensor decomposition, e.g. the Tucker model. In the decomposition of a Tucker model, the core tensor is not restricted to be super-diagonal and can have interacting values between multiple components across domains [42]. As an example, let $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ be a 3-way tensor and $\mathbf{A} \in \mathbb{R}^{I \times P}, \mathbf{B} \in \mathbb{R}^{J \times Q}, \mathbf{C} \in \mathbb{R}^{K \times R}$ be factor matrices for an *P,Q,R*-component Tucker decomposition. Then the Tucker-3 model is:

$$\mathcal{X} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} \mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r \tag{2.18}$$

where g_{pqr} is the $(p, q, r$-th) corresponding element of a core tensor $\mathcal{G} \in \mathbb{R}^{P \times Q \times R}$. Here, $\mathcal{G}$ defines the level of interaction between the different components [42], as illustrated in Figure 2.9. The CP decomposition of $\mathcal{X}$ is thus the same as the Tucker-3 decomposition when $\mathcal{G}$ is restricted to be super-diagonal, e.g. Figure 2.8 [42].

The flexibility given by $\mathcal{G}$ in Tucker, however, has a similar drawback to PCA and NNMF. Due to the flexible interactions between components across domains, additional constraints are

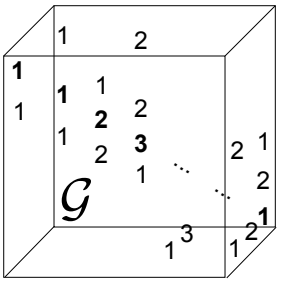

Figure 2.9: Pictoral representation of the flexible tensor core, $\mathcal{G}$ of a Tucker decomposition. The super-diagonal is bolded, with different number values representing various potential interactions across components and modes.

required to find a unique solution to the model. Furthermore, the flexibility in $\mathcal{G}$ compromises the ability to interpret how a component in one domain is reflective of changes in another domain. As the focus on this thesis aims to tie development reflected across components in the EEG of children, we have opted to use the CP model instead of Tucker, to exploit its guaranteed 1:1 super-diagonal core, mild uniqueness requirement and interpretable nature.

2.5.5 Component Selection in CPD/PARAFAC

Component selection is a critical step in tensor factorization [48]. The number of component factors chosen in a decomposition balances model suitability with proper representation of latent structural information. Appropriate choice for component number is especially critical in CPD/PARAFAC decompositions, as each mode must be modelled by the same number of components [42, 45]. Selecting too few components for CPD/PARAFAC may obscure the less dominant structural relationships across modes. Similarly, selecting too many components may lead to over-fitting in the model. The core consistency diagnostic (CORCONDIA) [48] was therefore used as a means to gauge the validity of specific *R*-component CP decompositions.

CORCONDIA provides an intuitive direct measure of how well a CP decomposition is able to model a tensor and thus can serve as a partial guide for model selection [48]. In brief, CORCONDIA measures the degree to which a given CP model deviates from its 'ideal' super-diagonal core [48], by comparing it to a the core tensor $\mathcal{G}$ from a Tucker decomposition [48]. To provide an example, again let $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ be a 3-way tensor and $\mathbf{A} \in \mathbb{R}^{I \times R}, \mathbf{B} \in \mathbb{R}^{J \times R}, \mathbf{C} \in \mathbb{R}^{K \times R}$ be factor matrices for an *R*-component CP decomposition. Given that the CP decomposition of $\mathcal{X}$ is considered as a special Tucker-3 decomposition where $\mathcal{G}$ is restricted to be super-diagonal [42], an approximate core tensor of $\mathcal{G}$ can be calculated by using **A**, **B**, and **C** [48, 49].

The deviation of the estimated core tensor $\mathcal{G}$ from the super-diagonal is then calculated as the percent CORCONDIA [48, 49]:

$$c = 100 * (1 - \frac{\sum_{i=1}^{R} \sum_{j=1}^{R} \sum_{k=1}^{R} (\mathcal{G}(i,j,k) - \mathcal{I}(i,j,k))^2}{R}) \tag{2.19}$$

where $\mathcal{I}$ is a super-diagonal tensor with zeroes everywhere except ones on the (r, r, r) entries

[49]. For a full description of CORCONDIA, please see [48].

The validity of a CP model can be estimated using the CORCONDIA value from Equation 2.19 [48]. A perfect modelling by CP with a pure super-diagonal core would give a CORCONDIA of 100% while flawed models would deviate significantly in their CORCONDIA value. Traditionally, models with CORCONDIA under 40% are considered to be non-viable. Low CORCONDIA indicates the number of components for decomposition (e.g. decomposition rank) may be inappropriate or that the underlying multi-linear structure may not be present [48, 49].

2.5.6 Evaluating Multi-way Component Outputs

Understanding the information gained from the multi-way analysis approach is crucial to support its use in a wide array of fields. Data gained from multi-way analysis of EEG data falls between directly measurable electrophysiological data (e.g. raw time-series EEG) and pure mathematical representations of the electrophysiological information (e.g. a strictly mathematical simulation of EEG). Therefore, there is a need to adequately understand exactly what the multi-way representation of EEG provides on a theoretical and practical basis. Together one can then understand how to interpret the information gained from the multi-way approach and, in the case of EEG, understand how it relates back to relevant brain physiology.

From the physiological point-of-view, using multi-way analysis to characterize EEG data provides a multi-relational representation of physiological events. As such, outputs provided by a multi-way model are not representing the raw physiological measurements themselves, but rather showcase the one-step removed relationship highlighting interactions occurring between these raw physiological measurements. Therefore, interpreting outputs from a multi-way model is done by evaluating the uncovered relationships between the raw data given to the multi-way model's basis, e.g. the measurable electrophysiological aspects of EEG data, and the relationships we may expect to see from such data, e.g. that the occipital region of the brain would likely be strongly related to activation occurring in the alpha frequency range.

To clarify, let us consider an example of using a 3-way CP model to represent rEEG with $[Spatial] \times [Spectral] \times [Temporal]$ modes. The raw EEG measurements for the $[Spatial]$ mode inform where the EEG activity was detected via scalp electrodes. The raw EEG

measurements in $[Spectral]$ domain inform on what the specific frequencies were present for the given rEEG time-series. Finally the raw EEG measurements used for the $[Temporal]$ mode informs on changes occurring over the rEEG recording interval. Together, these measurable physiological processess form the 3-way data array used in the CP model decomposition. The CP model then derives a set of factor matrices as an output representing specifically how these physiological measurements are interacting. That is, the CP model's output describes the specific relationship between events in one mode, e.g. the $[Spatial]$ mode, as they relate to events measured in the other modes, e.g. the $[Spectral]$ and $[Temporal]$ modes. Data given in the CP model's output for the $[Spatial]$ domain therefore reflect the relationship (e.g. contribution) from any given specific electrode as related to specific frequencies or temporal information rather than the direct measurement itself. Thus the multi-way approach provides a model depicting the multi-relational view of the interactions between the different raw, measurable electrophysiological data in the EEG. For further reading, please see [41, 43].

Through examining the underlying relationships and structure between aspects of data, multi-way analysis therefore provides additional useful information beyond traditional pair-wise investigations [38, 41, 43]. As the output from the multi-way analysis provides the multi-relational set of information, the relationships between each mode of data provided to the multi-way tensor are simultaneously evaluated across all modes without potential losses in structural relationships. Using pairwise investigations, e.g. PCA or NNMF, specific interactions between modes not included in the current pair of modes may be lost due to averaging over the data. To evaluate these lost relationships between the combinations of the pairwise data would then require additional steps, such as using a statistical evaluation approach to estimate how the pairs of modes are related. This is avoided in the multi-way analysis approach, as all of the structural information is preserved between the modes of interest. Furthermore, as the natural structure relating these modes are maintained, it is possible to evaluate how they change under specific perturbations from the multi-relational point of view in all the modes of interest. Finally, multi-way analysis provides a framework capable of handling a significantly large dimensionality in data. The multi-way tensor format views large dimensionality as a boon, with high dimensionality providing possibilities to obtain compact representations, to improve uniqueness of decompositions and improve flexibility in the choice of constraints [38].

2.6 Characteristic Filters and Direct Projection

In the CP decomposition, common interactions between different modes are encoded in the resulting factor matrices, due to the guaranteed 1:1 mapping from the CP's super-diagonal core. Hence it is possible to use combinations of the decomposed factor matrices to create 'characteristic filters' which can be used in the evaluation of new data. The decomposed weights from specific modes can be directly projected onto new incoming data to estimate its corresponding values in the modes of interest, without needing to re-run the CP decomposition [50]. This allows the characteristic filters to be effectively applied to previous data models, and offer a means to rapidly assess a whole new set of unseen data.

As an example, let $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ be a 3-way tensor. Let $\mathbf{A} \in \mathbb{R}^{I \times R}, \mathbf{B} \in \mathbb{R}^{J \times R}, \mathbf{C} \in \mathbb{R}^{K \times R}$ be the resulting *R*-component factor matrices of a CP decomposition of $\mathcal{X}$. Also, let $\mathcal{Y}$ be a tensor of similar design to $\mathcal{X}$, i.e. have the same mode designations in $\mathbf{A}', \mathbf{B}', \mathbf{C}'$. Then information from a given mode of $\mathcal{Y}$ (i.e. $\mathbf{C}'$) can be determined through direct projection onto $\mathcal{X}$ as follows:

$\mathcal{X}$ is unfolded (matricized) along the mode of interest (here, mode $\mathbf{C}$) using the Khatri-Rao product.

$$\mathbf{X}^{(3)} = \mathbf{C}[\mathbf{B} \odot \mathbf{A}]^T \tag{2.20}$$

or alternatively:

$$\mathbf{X}^{(3)} = \mathbf{C}\mathbf{W} \tag{2.21}$$

where $\mathbf{W}$ is an encoding matrix holding the estimated common interactions between modes $\mathbf{A}$ and $\mathbf{B}$. Unfolding the tensor in this way retains the structural relationship between the modes, while providing the encoding matrix needed to estimate values for new data held in $\mathcal{Y}$.

Estimating new weights for $\mathcal{Y}$ typically uses a least-square projection onto the encoding matrix, $\mathbf{W}$. First, $\mathcal{Y}$ is unfolded similar to $\mathcal{X}$:

$$\mathbf{Y}^{(3)} = \mathbf{C}'[\mathbf{B}' \odot \mathbf{A}']^T, \tag{2.22}$$

We then set $\mathbf{A}' = \mathbf{A}$ and $\mathbf{B}' = \mathbf{B}$ for projecting onto the known projections held in $\mathcal{X}$. Changing notation to match equation (2.21), we have:

$$\hat{\mathbf{C}} = \mathbf{Y}^{(3)}\mathbf{W}^{\dagger}, \tag{2.23}$$

where $\mathbf{Y}^{(3)}$ is the unfolded test tensor $\mathcal{Y}$ along the mode of interest, and $\hat{\mathbf{C}}$ holds the desired estimated weights for the new data.

One advantage of using these characteristic filters is that the properties of new data can be estimated rapidly for use in real-time applications. Tensor decompositions can become computationally expensive and memory intensive as more modes are included in the tensor analysis [51–53]. The low-rank component matrices derived from the multi-way decomposition, however, are far less resource demanding to store, and the direct projection outlined above requires only relatively simple matrix operations to find the estimated weights for the new data [50]. Characteristic filters and direct projection are used repeatedly throughout this thesis as a means to test a CP decomposition model on unseen data. Furthermore, as computation for the direct projection is only matrix multiplication, the characteristic filters could feasibly be used in real-time applications like BCI. In such cases, the proposed analysis would use a training session to first record the necessary EEG to construct the multi-way CP model of the data and identify the characteristic filters. Then, new EEG recordings could be evaluated via the direct projection in real-time through the estimated matrix multiplication. This would provide insight into the CP model's performance in the real-time task, as required by real-world applications of BCI.

2.7 Network Analysis and Functional Characterization of the Brain

Thus far, brain activity recorded via the EEG has been examined through a lens of structural connectivity, e.g. understanding how the physical neurons and structure of the brain relates to (and produces) the rhythmic oscillations measured by scalp electrodes and tools available to interrogate these structures. The presented multi-way analysis is shown to be a tool which can exploit this natural multi-way structure of the EEG to gain useful information

on brain processes [37]. Importantly, brain activity from the recorded EEG can also be examined from the viewpoint of functional connectivity. Functional connectivity aims to understand the brain processes by examining the inter-regional dependencies elicited from neurons' electrochemical impulses, thereby providing insight into how the brain's complex network of processes coordinates into functional outputs such as conscious thought, emotions, etc. [54, 55].

A variety of methodology has been developed to investigate functional connectivity in the brain. This thesis will focus on using network analysis of EEG to elucidate functional characteristics of the brain in particular. Network analysis can provide insight into how information flows across the brain, and how different neuronal activities may be collaborating [54–56]. Network analysis was selected as it is an established approach for characterizing functional brain connectivity via EEG [24,54–56]. Additionally, it provides a complementary set of information to the proposed multi-way analysis. To help support the complementary nature of the two methodologies, a brief introduction to the principles underlying network analysis is provided.

2.7.1 Basic Principles of Graph Theory

Network analysis of EEG data relies on the principles of graph theory, a branch of mathematics used to evaluate and model networks like the complex interaction between neurons in the brain [24, 54, 55]. The term graph in this context is not referring to plotting curves on an axis, but instead a specific mathematical object [57]. A graph is an object comprised of a set of points, commonly referred to as 'nodes', which share pairwise connections, typically drawn as lines, called edges [24, 54, 57]. Figure 2.10-A shows a basic example of a graph. Graph theory provides the tools used to interrogate and answer questions regarding properties of graphs, from describing the relationships between the edges and nodes to interpreting information present across the whole graph [54, 55, 57].

Network analysis of EEG utilizes graphs to decipher the functional connections in the brain and interpret brain processes [24, 54, 55]. As an example, let us consider two EEG signals signals *i,j*. Let signals *i,j* be recorded at any two electrodes, designated as nodes *i,j*. Then connectivity estimating the interaction between the pairwise signals of *i,j* at nodes *i,j* is represented as the relevant adjacent edge between the nodes, Edge *i,j*. The connectivity of Edge *i,j* therefore holds

the estimated weight between the relevant recorded EEG signals. Figure 2.10-B provides an illustration of this concept, for a general signal set *i,j* of interest. In this manner, the complex network of the brain measured is represented as a graph model object, using EEG electrodes as specific nodes, and estimated connectivity between these nodes as the edges of the graph. How connectivity between EEG electrodes (nodes) is estimated and derived is discussed in more depth in Chapter 7.

There are many metrics available for analyzing graph networks. The most fundamental network measure is the degree of a node, e.g. the number of connections (edges) that link a given node to the rest of the network [54]. Most other measures of a network are ultimately related to the node degree, such as finding the number of nodes in a network with degree 1, i.e. finding which nodes have only one connection. A node with one connection is commonly referred to as a 'leaf' node. The degree of all nodes in a network form a degree distribution, which can be Gaussian if all connections in a random network are equally probable. Alternatively, the degree distribution may be non-Gaussian in complex networks, like the brain. Assessing different degrees of nodes in a network reveals pertinent information on the topology of the network.

Another useful network measure is examining how neighbours of a given node are connected, e.g. the clustering present for a set of nodes. This is measured via the clustering coefficient, which quantifies the number of connections that exist between the nearest neighbours of a node (e.g. nodes within 1 edge) as a proportion of the maximum number of possible connections [54]. The distribution of different clustering properties provides information about the local interactions of a network and highlights what types of interactions a network can support.

Path length is another common network measure of interest in network analysis. Path length refers to the minimum number of edges required to go from one node to another. Complex and random networks tend to have short path lengths on average, indicating there is a high global efficiency in parallel information transfer across the network. More regular graphs would have longer path lengths indicating less global efficiency in parallel information transfer throughout the network.

Finally, the concept of hubs and centrality are important when analyzing graph networks.

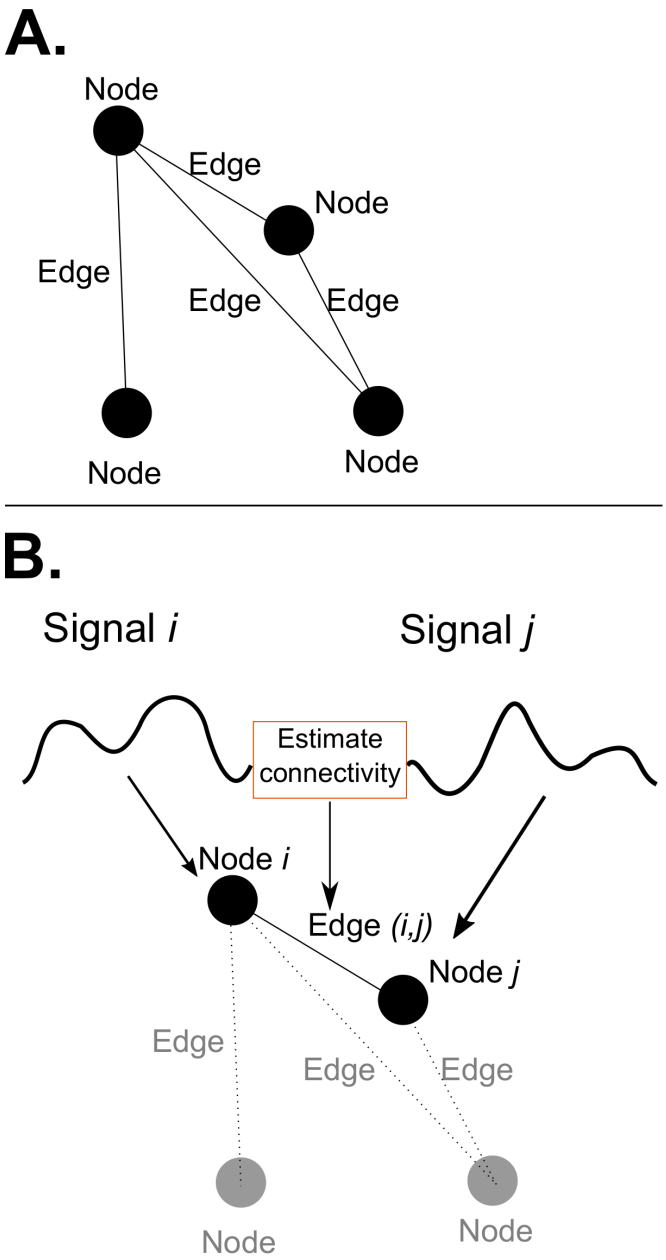

Figure 2.10: A.) Basic representation of an arbitrary mathematical graph object, with nodes and edges. Many types of graphs can be constructed, as discussed in Chapter 7. B.) Simple illustration showing the mapping of a graph from brain signals of interest, signals i,j, to the arbitrary graph given in A. Signals i,j are mapped to nodes i and j of the graph, with the connectivity estimation between these signals used as the weight of their respective edge, Edge (i,j).

Hubs are nodes with high node degree (e.g. >2 edge connections) or high centrality [54]. Centrality of a node measures how many of the shortest paths between all other node pairs in the network pass through the node of interest. If a node has high centrality, it is considered to be crucial when contemplating how communication would spread across the network. The higher centrality of a given node, the more critical it is as a hub for information flow across the network.

For more in-depth reading on graph measures and the basic principles of network analysis see [54]. A graphical depiction of several graph measures are given in Chapter 7 (see Figure 7.5).

2.7.2 Complementary Insights: Multi-way and Network Analysis

Network analysis and the discussed graph properties provide insight into how information and electric impulses throughout the brain may be organized during different activities [54, 55]. Through assessing the functional structure of the brain, network analysis is a useful tool in quantifying and interrogating information flow as it is processed across the brain. As such, insights gained from the network analysis are complimentary to multi-way analysis. For clarity, consider the following example scenario.

Suppose there is a given EEG recording from a child with epilepsy. The goal is to uncover pertinent information held in the EEG relevant to characterizing the child. As discussed in Section 2.5.6, outputs from the multi-way analysis will provide information on the structural relationships underlying the different EEG modes [37, 41, 43]. Thus, if the child is assessed at multiple recording points or during specific activities, the multi-way analysis will reveal the relationships between these different recording points and/or activities and alterations in the EEG's $[Spatial] \times [Spectral] \times [Temporal]$ modes. The insight gained from the multi-way analysis thus characterizes the child's data from a data structure perspective.

Now, let the recorded EEG be modelled as a graph and analyzed using network analysis. Information regarding the underlying *functional* brain network would be revealed [24, 54–56]. By using a set of graph measures mentioned in Section2.7.1, insights on the functional processing of the brain could be revealed. If the child is assessed at multiple recording points or during specific activities, the network analysis would reveal how information is flowing through

the brain and/or how functional connectivity is changed with different recording points and/or activities. Therefore, the insight gained from the network analysis characterizes the child's data from a functional network perspective.

Taken together, it is clear then that the multi-way and network analysis are providing complementary information. Both techniques are derive information from examining EEG, but provide different conditions and elucidate separate sets of information. The multi-way analysis provides a means to characterize higher-order relationships in the EEG data, while the network analysis delivers insight to the functional networks in the brain. The set of information gained can be used independently to answer separate questions regarding the characterization of the child's EEG. Thus, they can be interpreted as complementary approaches to characterizing EEG, where taken together they provide a more holistic set of information available from EEG recordings. This thesis examines the complementary nature of multi-way and network analysis in more depth in Chapter 7.

2.8 Matlab and Other Software Toolboxes

Several software toolboxes are used throughout this thesis in order to process data. This section briefly introduces the key toolboxes used in this thesis and their main role with respect to data processing and analysis.

2.8.1 Fieldtrip Toolbox

Fieldtrip is an open-source Matlab toolbox used for the analysis of EEG and other electrophysiological data, such as magnetoencephalography (MEG) [58]. Fieldtrip is fully implemented in Matlab through nearly 1000 high-level and low-level functions, available for users to adapt to various electrophysiological scenarios. Fieldtrip's functionality includes algorithms for data pre-processing, event-related response, parametric and nonparametric spectral analysis, source modelling, connectivity analysis, classification, real-time data processing, statistical inference and data visualization. A variety of options are available for each of these topics, providing a deep set of tools available for researchers.

The main focus of the toolbox is the analysis of electrophysiological data but Fieldtrop can

be used for processing any time series data in theory, e.g. functional magnetic resonance imaging (fMRI) or functional near-infrared spectroscopy (fNIRS) recordings [58]. The toolbox supports reading a large number of different file formats, ranging from manufacturer specific files to open source file types. To this end, Fieldtrip aims to be a common platform for both experimental scientists and developers at all levels.

Fieldtrip is used extensively throughout this thesis for many of its functions. In particular, pre-processing and spectral analysis are handled by Fieldtrip for most datasets presented. Connectivity analysis (as part of the network analysis methodology in Chapter 7) is also handled by Fieldtrip. As such, this toolbox was key in developing the analytic codes used for this thesis. Please see the associated Fieldtrip documentation (`http://www.fieldtriptoolbox.org/documentation/`) for a more in-depth description of the available Fieldtrip toolbox functions.

2.8.2 N-way and Tensorlab Toolboxes

The multi-way analysis in this thesis relies on two Matlab toolboxes: The N-way toolbox (Chapter 4, 5) and the Tensorlab toolbox (Chapter 5, 6). The N-way toolbox is a freely available collection of functions and algorithms specifically designed for modelling multi-way data sets using multi-linear models, put together by C. Andersson and R. Bro [59]. The N-way toolbox provides several multi-way model algorithms directly, such as the CP model, as well the functions required to design new multi-way processing methods. Furthermore, the N-way toolbox provides a number of constraints available for analysis, including non-negativity and unimodality. This thesis takes advantage of the N-way toolbox's developed CP model and associated unimodal constraint primarily for the analysis of rEEG in Chapter 4. Please see the associated documentation (`http://www.models.life.ku.dk/nwaytoolbox`) for additional information on the available multi-way functions from the N-way toolbox.

Another multi-way analysis toolbox for Matlab is the Tensorlab toolbox. Tensorlab also provides a wide variety of functions built for multi-way tensor computations [60]. Of note, Tensorlab includes built-in functions suitable for coupled tensor-tensor and tensor-matrix decompositions, via structured data fusion [60]. Additionally, Tensorlab provides direct support for investigating specific types of multi-way tensors, e.g. incomplete, sparse or

structured tensors, not readily available in the N-way toolbox. However, only a limited number of constraints are directly available in the Tensorlab toolbox. Tensorlab's built-in functionality for data fusion is exploited in Chapter 6 of the thesis. Please see Tensorlab's associated documentation (`https://www.tensorlab.net/userguide3.pdf`) for more information regarding the available multi-way functions.

2.8.3 WEKA Toolbox

The Waikato Environment for Knowledge Analysis (WEKA) Data Mining Software is a freely available workbench for state-of-the-art techniques in machine learning [61]. The WEKA project aims to provide a comprehensive collection of machine learning algorithms and data processing tools to researchers and practitioners alike. The WEKA toolbox is open-source and modular, providing an extensive library of sophisticated machine learning algorithms and processes developed by the research community. A base set of machine learning algorithms is included in WEKA, allowing classification and regression schemes to be specifically tailored to a given question. WEKA is used in this thesis for the classification schemes developed in Chapter 4, 7. The open architecture of WEKA provided a basis of useful tools for the machine learning aspects of the thesis. For more information, please see WEKA's documentation (`https://www.cs.waikato.ac.nz/ml/weka/documentation.html`)

2.9 Conclusions

This chapter provided background information necessary to understanding key multi-disciplinary topics fundamental to the thesis. Common methodology used throughout the thesis was described, and general examples were provided for essential concepts like the CP model decomposition and network analysis. Relevant neurophysiological context for the thesis was also covered. A deeper examination of the motivation for this work and its association to BCI applications in children is covered extensively in Chapter 3.

Chapter 3

Reviewing an Unmet Need: EEG-based BCI for Motor Rehabilitation in Children

The emerging technology of brain-computer interfaces (BCIs), and in particular rehabilitation applications of BCI, served as the real-world application which inspired the beginning months of this doctoral thesis. At the time, there was quite a gap present between reported BCI research and its realization for children. This gap was especially prevalent with respect to motor rehabilitation based BCI applications for children, where there was no clear evidence indicating why only limited research had been pursued in this application. To address this unmet need in an organized manner a formal review of the literature surrounding BCI motor rehabilitation in children was produced.

This chapter thus presents the formal review of the BCI field for children, published in the *Journal of Neural Engineering* in 2016 [1]. The literature review serves as a frame of reference, outlining the real-world motivation which inspired the technical advancements that were pursued in characterizing paediatric EEG data, as described in Chapters 4-7 of this thesis.

Due to its nature as a review, the structure in this chapter reflects the literature present in 2015 during the inception period of the doctoral research. As such, it is largely presented in its original state, i.e. focusing on establishing literature support for advancements we ultimately present later in the thesis. Since its initial publication, there has been an increase in research activity aiming to bridge the gap between BCI applications and its translation to children. Therefore, an updated section looking at recent literature contributions inspecting BCI applications for children is included at the end of this chapter.

The main contributions of this chapter include:

- A thorough review of BCI applications for children.
- An in-depth examination of the unmet need that EEG-based BCI could offer for rehabilitation in children.
- Justification and support for improving methodology and advancements in EEG-based BCI and its importance in translating such technology to children.

3.1 Introduction

The past decade has seen a significant increase in researchers pushing the limits of BCIs, devices capable of creating a non-muscular communication channel directly between the brain and an output source, such as a computer [8, 62]. BCI has emerged as a malleable technology, viable in a myriad of applications (for a general review see [8]). Historically, the non-muscular nature of BCI lent itself to use in communication applications and assistive technologies for patients suffering from conditions such as amyotrophic lateral sclerosis (ALS) or persistent locked in state (LIS) (for reviews on these topics see [63–65]).

Rehabilitation applications using BCI have recently shown encouraging results for motor recovery, harnessing the underlying properties of brain plasticity for rehabilitation [66, 67]. Rehabilitation focused BCI applications range from neurofeedback therapy for autism spectrum disorder (ASD) [68], and attention deficit hyperactivity disorder (ADHD) [69], to schizophrenia [70–72] and motor rehabilitation post injury, such as after a stroke [67, 73–75].

Current motor rehabilitation techniques, however, can be limiting to some patients as they require residual movement in afflicted appendages and are regarded as potentially too demanding [63, 76]. Non-invasive motor rehabilitation BCI offers an alternative therapeutic approach accessible to these individuals [63, 76].

Despite the expanding breadth of current BCI research, the majority of investigations have focused on BCI applications for adults [77]. This is especially evident in motor rehabilitation applications of BCI [77]. Such a restriction undercuts the possible benefits that BCI technology could bring to children who suffer from neurological diseases and disorders. The prevalence of motor impairment from neurological impairments such as cerebral palsy (up to 4 children

per 1000 are diagnosed annually with cerebral palsy according to the Centers for Disease Control (CDC) [78]), paediatric stroke (approximately 11 in 100,000 children diagnosed per year [79]), and traumatic brain injury (approximately 3,000 children aged 0-14 will have a TBI per 100,000 in the United States [80]), among other conditions, present a substantial population which could benefit from the recent advances in BCI technology. Pairing recent breakthroughs in understanding the role of brain plasticity in successful therapeutic BCI applications [66, 67, 74, 81, 82] with considerations on the abundant plasticity in the paediatric brain [83, 84] and the (relatively) limited time since injury, developing early-life BCI applications for motor rehabilitation becomes well-posed as a valuable, unrealized research pursuit.

A systematic search was done to explore the state of research concerning motor rehabilitation using BCI for children. Surprisingly, the systematic search into the literature reflected an almost barren field of research with respect to EEG based BCI motor rehabilitation in children. Further, there was limited research on any published BCI rehabilitation applications incorporating children. This lead to the following questions:

1. Do trends in the literature support developing EEG-based BCI motor rehabilitation applications for children?
2. Is there clear evidence in related investigations which identify and implicate factors that would hinder the development of such applications?

To answer these questions, literature surrounding topics of EEG-based BCI, motor rehabilitation applications, and successful BCI for children are examined. Evidence provided by this 'indirect literature' helps provide a clearer picture of the feasibility of EEG-based BCI for motor rehabilitation of children, despite the lack of explicit research evidence. As such, this is not an exhaustive review of each topic related to EEG-based BCI, BCI rehabilitation, rehabilitation in children, etc. There is an abundance of such reviews examining those subjects[1]. Instead, this work investigates the above questions using literature evidence from a diverse, exploratory set of reports in hopes to shed light on the feasibility of EEG-based BCI

[1]Examples of exhaustive reviews on several related topics- Neurofeedback rehabilitation with children is covered in [69]; BCI in rehabilitation and its current status post stroke is reviewed in [67]; A look at assistive technology integration with BCI is detailed in [63]; Electrocorticography (ECoG) and implantable BCI used in motor rehabilitation is examined by [85]; Neural plasticity and its role in BCI applications is surveyed in [73]; BCI design and interpretation pitfalls to avoid are discussed in [86]; Current physiotherapy applications are considered in [87].

motor rehabilitation in children.

The chapter is organized such that:

- In Section 3.2, the systematic search conducted on the Pubmed.org database, and whose results helped define the indirect categories used for establishing evidence in this review, is described in detail.
- Then in Section 3.3, a current look at investigations which have successfully implemented a BCI motor rehabilitation paradigm for adult patients is discussed. Considerations are given to how this literature supports/dissuades motivation for similar applications in children.
- Following in Section 3.4, EEG-based BCI paradigms which have included child populations are examined, with particular examples and roadblocks associated with developing such BCI applications highlighted.
- Afterwards in Section 3.5, recording modalities outside of EEG, specifically electrocorticography (ECoG), is briefly presented alongside what achievement in alternative modalities implies for EEG-based BCI.
- Then in Section 3.6, justification for more in depth examination of BCI in children is given, with emphasis on how to address potential solutions to the translational barriers presented in the previous sections. Trends in the literature are synthesized to answer the initial questions regarding the potential of EEG-based BCI motor rehabilitation in children.
- In Section 3.7, an update on new publications in the field since the original review is covered. Special attention is given to how research into BCI for children has progressed since the initial review and call for increased attention to this unmet need.
- Finally in Section 3.8, limitations and conclusions from the literature survey are provided.

3.2 The Systematic Search

A systematic search of the Pubmed.org database was conducted on May 04, 2016. The Pubmed database contains more than 27.3 million records dating back to 1966, and is

maintained by The United States National Library of Medicine, at the National Institute of Health (NIH). The search included a combination of multiple key words and related synonyms including 'children', 'kids','paediatrics', 'BCI', 'Brain-computer interface', 'Brain-machine interface','BMI' and 'rehabilitation'. Results from all combinations of key terms were saved, and examined for relevance with respect to BCI rehabilitation in children. Inclusion or exclusion of literature was done in two phases. First, manuscripts were excluded based on titles definitively not related to rehabilitation in children through BCI and/or neurofeedback means (i.e. 'Spatial knowledge of children with spina bifida in a virtual large-scale space'). First pass inclusion criteria was lenient to avoid possible exclusion of related manuscripts. Abstracts of the remaining publications were then examined, with those directly related to BCI then included. All remaining articles were then examined in depth, with relevant references investigated for additional insight and information. Results were categorized based on major themes present, and separated accordingly. In our inclusion criteria, and therefore throughout this chapter, individuals between the ages of 2 and 16 years old were considered as children. Figure 3.1 illustrates a visual breakdown of the search, along with the categorized results.

It is important to note that the systematic search revealed no study explicitly examining EEG for motor rehabilitation through motor imagery (MI) for children. There was a single manuscript, [88], which mentions one of its participants was aged 12 while another three were age 16. However, the manuscript focuses on BCI used with assistive robots rather than neurorehabilitation. Therefore, it is reflected as an '≈ 1' instead of a hard '$= 1$' in the breakdown of the systematic search in Figure 3.1.

The category of virtual reality (VR) literature was considered to be a grey area with respect to what is considered as true BCI applications. Although VR studies can have aspects of BCI, such as haptic feedback coupled with a virtual environment stimulus, interaction with the VR was not necessarily accompanied or driven by direct recordings of the brain (i.e. [82, 89]). Therefore, we have elected to keep the VR literature separate from our consideration of BCI. However, considering its related nature we have not excluded it completely to the 'Out of Scope' category. We refer interested parties to the reviews [90, 91].

Finally, it is important to mention commercially available BCI toys and games (i.e. [92–95]). These technologies were excluded from our systematic review for several reasons. First, they

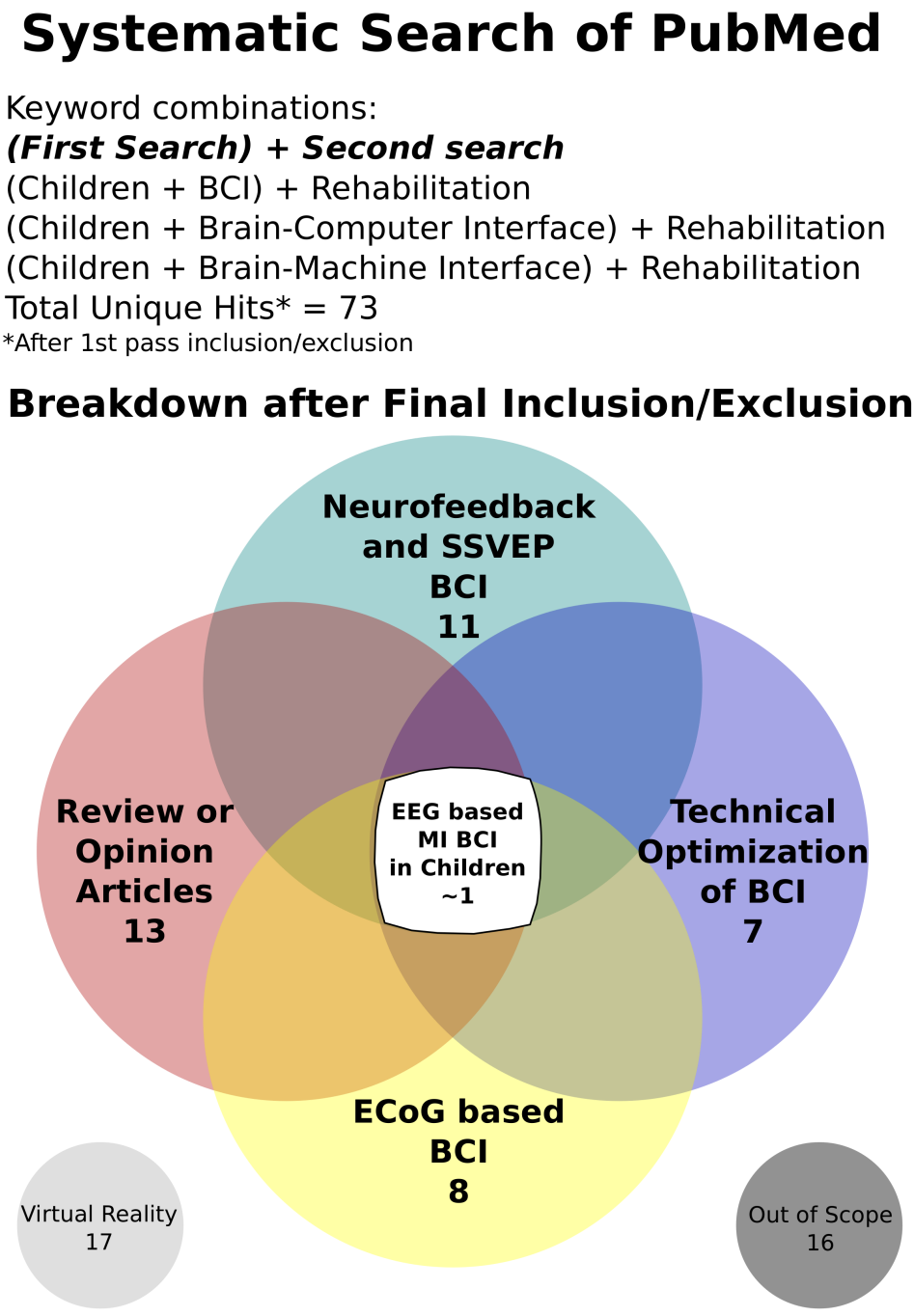

*Figure 3.1: A visual representation of the results from the systematic search of Pubmed.org. Acronyms: **SSVEP**- steady-state visual evoked potential; **MI** - motor imagery.*

are not documented as related to neurorehabilitation or motor rehabilitation and were not found during our systematic search. Additionally, their intended purpose is out of scope for the present review as it is not rehabilitation focused. Further, the structure of these toys often relies on a minimalistic EEG set-up, basing their BCI signal acquisition on only a few electrodes. These minimal electrodes do not often include covering the whole of the motor cortices, but are rather placed on the frontal or temporal lobes [92–95]. The driving signal from these tend not to be from MI, which is the control paradigm often used in BCI motor rehabilitation [67]. To the best of our knowledge, excluding these commercially available products does not affect our analysis of the current state of BCI for children. However, the relative success of these toys suggests that children are likely to find BCI technology attractive. Perhaps, then, children would find BCI more engaging for rehabilitation than other alternatives, such as constraint movement rehabilitation, thus supporting the need to fully explore the potential of EEG based motor rehabilitation BCI applications for children.

3.3 BCI Therapy: Motor-imagery BCI in Rehabilitation Applications

3.3.1 MI-BCI and Neurorehabilitation: Initiating Plasticity Through Thoughts in Post-stroke Patients

A recent BCI application generating interest is the use of motor imagery BCI (MI-BCI) in therapeutic paradigms to improve impaired motor control due to trauma or disease, such as stroke (for a general review see [67]). MI-BCI functions through modulating sensorimotor rhythms over the primary somatosensory and motor cortical areas through imagined self-movement, which causes a measurable event related desynchronization (ERD) to occur [67], [81]. Imagined movement generates other measurable potentials that can be integrated with BCI technology, like the movement-related cortical potential (MRCP) associated with pre-motor planning of volitional movement [74]. Literature evidence shows that MI and MRCP-based BCI alike can induce pathways which influence the underlying mechanisms responsible for brain plasticity [67, 74, 81]. One concrete example showing this relationship is from the work of Pichiorri *et al.* [81]. The authors demonstrated that

successful MI-BCI lead to increased motor cortical excitability, visible within 24-48 hours post intervention in ten healthy subjects [81]. The authors suggested that goal-oriented MI-BCI training could re-instantiate more efficient connections within motor cortical areas, leading to potentially better motor recovery [81].

The suggestion for more goal-oriented MI-BCI training has been adapted, to a degree, by pairing BCI with stimuli to enhance possible Hebbian-associated learning. In a proof-of-concept study by Mrachacz-Kersting *et al.* [74], a transcranial magnetic stimulus (TMS) was used to measure how Hebbian-associated feedback with MRCP-BCI potentially improved lower-limb function. A subset of 22 chronic-stroke patients received Hebbian-associated feedback with MRCP-BCI while no Hebbian-associated feedback was given to the another subset. The authors report a significant increase in the power of the associated tibialis anterior (TA) muscle evoked potential in the associated group that was absent in the non-associated group [74]. The authors highlight that the associated learning method employed in their work may be crucial to unlocking the full potential of BCI induced neuroplasticity for motor rehabilitation [74].

Other researchers have also reported positive changes in stroke patients from MI-BCI for motor rehabilitation [76, 82]. In a pilot study, Young *et al.* [82] successfully exploited the reorganization function of brain plasticity to promote functional improvement in upper-extremity movement for eleven post-stroke victims suffering from motor loss. Using a closed loop, multi-modal neurofeedback (NFB) BCI which included visual, functional electric stimulation (FES) and tongue stimulation, the group discovered positively correlated gains in both objective (neural responses to BCI treatments) and subjective (self-reports on improvement after treatment) measures after MI-BCI intervention [82]. The investigators used the laterality index (LI) of functional magnetic resonance imaging (fMRI) to measure the functional brain organization of patients before and after MI-BCI therapy [82]. They found a reorganization and re-lateralization of the lesioned cortex in response to MI-BCI therapy which corresponded to improved motor recovery by patients [82]. This finding echoed results described in a case study by Caria *et al.* [76]. Caria and company explored the motor recovery of a chronic stroke patient who had undergone combined BCI training and physiotherapy [76]. The authors report functional changes in the brain organization,

with a significant re-lateralization towards the perilesional cortex post combined BCI training measured by several neuroimaging modalities [76]. Interestingly, physiotherapy alone was not enough to induce such neuroplastic alterations to the laterality index and the re-organization of structural white matter [76]. Changes were only found after the combined BCI-physiotherapy training [76].

Throughout these studies the theme of improved motor function, due to MI or MRCP BCI enhanced therapy causing activation of neuroplastic pathways, illustrates a trend of positive improvements of BCI applications for motor rehabilitation (some additional examples include [75, 89, 96]). The studies mentioned are in no way an exhaustive list of studies looking at BCI applications for stroke, nor are they intended to cover all approaches within this broad research field. They were selected to help highlight the use of MI-BCI as a platform for inducing neuroplasticity. These studies fulfill a call for more empirical evidence demonstrating MI-BCI as a neurorehabilitation tool [73], and provide insight into addressing the initial proposed questions with respect to motor rehabilitation BCI applications in children.

3.3.2 BCI Control and Motor Rehabilitation: Pathologies Outside of Stroke

Displaying a functional control over MI based BCI implicitly provides a basis for BCI motor rehabilitation applications due to MI-BCI inducing activity-dependent neuroplasticity [66, 82]. Exploring control of MI-BCI in populations beyond stroke can help support development of BCI motor rehabilitation applications for patients across the spectrum of motor disability. For example, Faller *et al.* [97], demonstrated users with severe motor impairment including multiple sclerosis (MS), traumatic brain injury (TBI) and spinal cord injury (SCI) could use MI-BCI. Through the use of a novel co-adaptive training paradigm, which provided immediate feedback of the user's brain-activity with a continuously updated underlying classifier model, 22 users were able to effectively control a MI-BCI [97]. The authors suggest the co-adaptive BCI training paradigm could be a potentially useful tool in neurorehabilitation for future work [97]. In a similar vein, Daly *et al.* [98] explored the practicality of BCI control in 14 cerebral palsy (CP) patients naive to BCI using several different control schemes, including sensorimotor rhythm modulation, like MI. The study reports at least six of the subjects could control MI-BCI to a significant degree, despite only interacting with the BCI system for a relatively limited period of time over the course of a few days [98]. The intent of the paper was

to demonstrate the feasibility of control, and thus extended training and interaction with the BCI system, which may improve subject control over the system (as seen in [99] [42]), was not examined [98]. Why only some users possess innate abilities to use BCI without prior exposure or training is an ongoing topic of research and an open question (for a review see [100], for other examples see [101, 102]). Further, a pilot study by Cincotti *et al.* investigated control of assistive technologies in 14 able-bodied subjects, and 14 patients suffering from Spinal Muscular Atrophy type II (SMA II) or Duchenne Muscular Dystrophy (DMD) through both standard muscle input and MI-BCI [88]. Their goal was to evaluate the effectiveness of BCI and standard inputs for using commonplace technology (i.e., accessing a computer cursor) and aide technology (i.e., as a controller for intelligent motion devices) [88]. The researchers found when using a BCI application, able-bodied subjects were able to control a standard application of the BCI (i.e., moving a cursor on a screen) with an overall accuracy above 70% [88]. Retention of control was substantially maintained when using BCI to drive other environmental output devices as well [88]. The authors report similar levels of performance were achieved by the four motor impaired patients who underwent BCI training [88]. (Although the authors do not explicitly mention which patients participated in the standard BCI training, it is important to keep in mind that the study included three patients aged 16 or younger who were possibly included.) These results illustrate BCI applications can be controlled with reasonable accuracy for both DMD and SMA motor disorders [88].

These investigations highlight that MI-BCI can be controlled by a diverse user population across the motor impaired spectrum. They also illustrate these groups as potential candidates for MI-BCI motor rehabilitation applications. It is important to note that although rehabilitation applications have been highlighted, more overtly assistive technology (e.g., a BCI spelling machine) may be more beneficial for some patients with motor impairment, especially in extreme cases such as LIS. Decisions to more heavily stress assistive, rehabilitative or a combination of BCIs should be assessed at an individual patient level, taking into consideration factors like residual movement and tolerance to available technologies [88]. Additionally, rehabilitation applications may be considered as implicitly assistive, blurring the decision on which BCI paradigm to emphasize [103]. The studies mentioned here have been selected to highlight the interaction between control of a MI-BCI and its related rehabilitation or assistive benefits in several different adult motor impairment pathologies. These examples illustrate

a key concept needed in translating such applications to the developing brain and its varied pathologies.

3.3.3 Neuroplasticity and Age: Motivation for Translating MI-BCI Motor Rehabilitation to the Developing Brain

The positive literature trends in neuroplastic initiation via BCI supports motivation for developing motor rehabilitation BCI for children. Core to general BCI rehabilitation applications is the interaction between patient controlled BCI, the corresponding activation of neural plasticity and the correlated functional improvement in the patient. In their publication, van Dokkum *et al.* stresses the importance of properly selecting time windows for optimal BCI neurorehabilitation, asking the question of when BCI could be applied to gain the most from plasticity [67]. Looking at this question broadly, children may be ideal candidates for therapeutic motor rehabilitation using MI-BCI intervention. For example, they are ideal due to the enhanced natural plasticity of the developing brain, evidenced by its remarkable ability to recover from early injuries and enhanced capability in learning [83, 84]. Earlier therapeutic intervention via BCI applications could also have additional benefits when considering temporal factors like regional recruitment (the brain's capability to recruit new pathways to compensate for deficits) and learned non-use of afflicted regions post injury [84]. This sentiment is echoed by Daly *et al.* [104]. In discussing their results, the investigators report a lower event-related desynchronization (ERD) strength in patients with lesions occurring in early childhood, as compared to patients who had lesions occur in adulthood (e.g. stroke) when using MI-BCI [104]. They hypothesize that recruitment of the cortical areas for other functions or slower learning of motor processes, and by extension possibly motor imagery, may explain the relative difference in impaired ERD strength for the 14 cerebral palsy patients in the study [104–107]. The authors suggest that introduction to BCI rehabilitation early in childhood is a promising route which could encourage greater ERD strengths for patients and potentially lead to improved usability of BCI applications [104]. This could impact sustained BCI rehabilitation over time, as familiarity with the technology reduces frustration, leading to increased use and potentially better long-term improvements [99, 108].

Earlier exposure to BCI induced therapy could also potentially improve the re-lateralization of

the brain, which then elicits greater improvements for traditional therapeutic approaches [109]. In a study by Manning *et al.* [109], researchers found a positive neuroplasticity activation in seven children with cerebral palsy using the relatively common therapy of constraint-induced movement for motor recovery. Crucially, Manning *et al.* mention that the greatest improvement in motor recovery was found in children with the greatest asymmetrical laterality indices at the start of the intervention [109]. They also report the inverse situation to be true; namely, the least improvement corresponded to children with a brain baseline which was highly bilaterally organized [109]. These findings are possibly the result of recruitment of those regions for other functions [109]. Therefore, early-life therapeutic intervention could lead to a more effective rehabilitation for BCI. Changes in a child's brain plasticity resulting from therapeutic interventions could be monitored through cutting-edge tools and neuroimaging modalities, like functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI), to assess efficacy [76, 110].

Motivation for developing BCI based motor rehabilitation applications for children is clearly supported by the evidence in current adult literature. Also, almost no evidence of an explicit barrier impeding development of such applications was discovered in reviewing this set of literature. Improvement concerns for BCI in general were present, but nothing explicitly illustrating why there has been no attempt at such a BCI paradigm. Although barriers may well exist, we aim to investigate this concern in the next section using the literature evidence from groups who have implemented other BCI paradigms in children. The literature examined is the indirect results obtained from our systematic search.

3.4 Barriers to Entry: Neurofeedback Applications and Designing BCI for Children

3.4.1 Growing into BCI: Current Literature Results From the Systematic Search

The systematic search revealed a sparse set of studies actually using BCI applications to treat children, a trend observed previously [77, 111, 112]. Manuscripts which explicitly report using EEG based BCI paradigms are highlighted [113–115] here in hopes to elucidate current

approaches to BCI in children and potential pitfalls which may inhibit translating motor rehabilitation BCI applications to children.

Research concerning child-focused BCI applications is limited in terms of both empirical examples and success. While the majority of the literature describing BCI applications with children have focused on neurofeedback applications aimed towards developmental disorders such as attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD), few groups have successfully implemented these paradigms (for a general review see [69]). One group, Lim *et al.*, successfully developed an attention-based BCI training program geared towards improving ADHD symptoms on 20 unmedicated children aged 6-12 years old [113]. Their attention-based BCI-training program included an engaging, 3D graphic game as the user interface for the therapeutic BCI intervention, with 8 weeks of intervention [113]. The authors remark that with BCI intervention, parent-reported symptoms of ADHD were significantly improved with sustained results evident as far out as 3 months post treatment [113]. Another study by Rohani *et al.* describes a prototype BCI system which uses the P300 potential as the driving feedback in two separate VR games [114]. The investigators developed the set of VR games to reward simultaneous stimulus of the P300 with correct information gathering as a means to measure subject attention for five healthy young subjects [114]. The authors conclude that the preliminary data demonstrates that a system using the P300 potential can be used to measure attentiveness of subjects [114]. They also suggest an approach to game design with respect to addressing distractions for ADHD individuals [114]. Outside of these two neurofeedback manuscripts, only one other group explicitly mention using an EEG based BCI with children. In their paper, Ehlers *et al.* investigated the extent development-specific changes in background EEG would influence the ability of children age 6-10 years to control a stimulus-driven steady-state visual evoked potential (SSVEP) BCI [115]. The researchers investigated a total of 14 adults and 37 children between age 6 and 33 years [115]. The children were separated into approximately 3 equal groups, resulting in a mean age for each group of 6.73, 8.08 and 9.86 [115]. Their results report a significant difference in ability to control a SSVEP-BCI spelling application between the children and adults, with greatest disparity between the youngest group (approximately age 7) and the adults [115]. Ultimately, they highlight that an appropriate SSVEP-based BCI system for children currently remains purely academic [115].

Although not directly empirical, Friedrich *et al.* [68] proposed a combined neurofeedback and biofeedback treatment for children on the autism spectrum through a BCI game application. The authors underscore the importance of entrenching the positive benefits of BCI therapeutic intervention, in this case via neurofeedback, in a medium of play for children [68]. Friedrich and company speculate that play is an ideal medium to encourage and engage children in sustained interactions, especially in BCI design for children on the autism spectrum [68]. This perspective on play may be a critical philosophy needed when considering developing any BCI applications for children.

3.4.2 Potential Barriers to BCI in Children: Limitations in the Literature

With respect to the literature examined above, there were several limitations and concerns which should be kept in mind when interpreting and expanding upon the results. In the study by Lim *et al.* the authors highlight that the intervention was well tolerated, but a side effect of a mild headache was reported for two of the participants [113]. This side effect, however, did not stop them from continuing with the treatment [113]. Additionally the uncontrolled, unblinded nature of the study may have a biased treatment effect [113]. In the work by Rohani *et al.* the system may need to be redesigned to be useful in a rehabilitation setting through improving the error rate and reliability of feedback to the user [114]. Additionally, concerns with the reliability of a P300 as a measure of attention should be addressed as the user improves their attention [114]. Potential design limitations are present in the work from Ehlers *et al.* as well. The authors mention that the observed age-group differences might be due to development-specific deficits in dealing with visual search tasks for children under ten years [115]. Considering the SSVEP-BCI spelling system relies on the ability to spell, the 79% failure rate for completing tasks by the youngest age group, approximately 7 years, may be partially explained as a developmental limitation in the design paradigm [115]. This concept is further supported by the decreased failure rate corresponding to increasing age [115]. These limitations do not encompass all possible concerns with translating BCI to children, but help serve as guides when drawing conclusions from the current state of BCI applications which include children. A summary of these studies is provided below in Section 6.2, Tables 3.1, 3.2, 3.3.

3.5 Alternative Inputs for BCI Control: A Glance at ECoG

3.5.1 ECoG and EEG

Beyond the scope of EEG-BCI lies a handful of studies using alternate inputs to directly investigate motor imagery based BCIs in children. Again, these indirect studies were discovered through the systematic search and examined for evidence to shed light on the feasibility of EEG based motor rehabilitation BCI for children. A variety of input recording technologies have been used to drive BCIs, including magnetoencephalography (MEG), electrocorticography (ECoG) and electroencephalography (EEG). The non-invasive, cheap and portable advantages of EEG based BCI (EEG-BCI) provide certain benefits for developing BCI applications, leading to it becoming the most widespread recording modality [8, 30]. The relatively cheap components and portable nature of EEG-BCI means a larger number of potential patients and more convenient use [30]. The trade-off for these benefits is the relatively poor spatial resolution of EEG [8, 30]. Since EEG-BCI exploits changes in electric currents on the scalp, there is higher noise in recorded EEG signals due to the signals crossing through the skull and scalp before reaching the EEG electrodes [30]. Further, the high noise and lower power in EEG may inhibit higher frequency signals (> 40 Hz) from being reliably deciphered and interpreted [16, 30]. These disadvantages are not present in other input modalities, like ECoG. ECoG offers a greater spatial resolution, higher sensitivity to neural activity and larger discernible electro-physiological spectrum than EEG [30]. The main trade-off for the benefits in ECoG is the highly invasive craniotomy procedure required to place electrodes on the cortical surface [30]. Due to how invasive this procedure is, ECoG based BCI have often been implemented only with patients already scheduled for surgery [116]. Therefore, current ECoG based BCI is an option for only a small subset of the population.

3.5.2 Paediatric BCI Studies Using ECoG

The systematic search revealed several groups successfully used ECoG-based MI-BCI with children. In a pioneer study by Breshears *et al.* [116], researchers revealed that signals from the paediatric cortex could successfully be decoded to control a MI-BCI to an accuracy similar

to those found in adults. Six paediatric patients aged 9-15 were able to rapidly achieve control over the BCI using both overt and imagined cognitive modalities, including using MI to control a game or an external robotic hand [116]. Importantly, the results exhibited no significant differences between the time or accuracy needed for the paediatric patients to gain control of the BCI compared to adults [116].

The authors promote these results as a successful proof-of-concept that decoding signals from the paediatric cortex is possible for use with motor imagery based BCI [116]. The capability to decode, identify and classify signals from the paediatric brain for use in a BCI, in spite of significant electro-physiological differences, is a promising finding for translating BCI technology to the paediatric brain. It is important to note that the majority of features used to control the BCI here were attained from high-gamma bands (60-130 Hz), a frequency range in EEG which presents difficulties in parsing signals of interest from noise [30, 116]. However, improvements in signal processing and artefact detection could potentially unlock these high-gamma bands for use in EEG [14]. At a minimum, four of the six participants were able to control the BCI with signals from the primary motor and pre-motor areas using the classic beta band (15-40 Hz in this paper) as well [116]. Additionally, the researchers Roland *et al.* [117] investigated the effect of age on ECoG signals and evaluated their implications for BCI applications for 23 patients aged 11 to 59 years. The researchers found that the magnitude of percent change in power for all low-frequency bands was not largely correlated with age, while high-frequency bands showed significant correlation [117]. However, a correlation between age and area of activation for the alpha/beta bands and a correlation between age and cortical networks for just beta bands was found [117]. Roland and company conclude that the more stable the signal platform is, and by extension the underlying physiology, the more likely control of the BCI will remain reliable with time [117]. Thus they argue for the use of high gamma rhythms as the best choice for long-term ECoG-BCI use [117].

The results from these ECoG-BCI studies demonstrate strong BCI control through modulating motor imagery in children is possible. This supports the conclusion that motor rehabilitation BCI could be functional for children, at least with respect to using ECoG. When generalizing to our main question of using EEG-based BCI for motor rehabilitation, some barriers present in these ECoG studies provide additional considerations for development. With respect to long

term motor rehabilitation, the issues arising due to differences in changing cortical maturation of children is an important concept. Additionally, advancing signal processing techniques to mimic the improved signal sensitivity in ECoG compared to EEG could help further realize an EEG-based motor rehabilitation BCI for children.

3.6 A New Hope: Justification for BCI and Prospective Solutions

3.6.1 Justification for BCI in children

BCI technology has innate advantages for use in neurorehabilitation which can be further capitalized upon during early-life intervention. The main advantage of BCI is its accessibility to all levels of physical deficits. Current physical therapy (PT) treatments, like constraint-induced movement (CIM), may not be possible for patients when residual movement and control is below certain thresholds, as mentioned in [63, 76]. BCI, on the other hand, is accessible to all individuals with physical deficits despite their level of residual control. Since early-life motor rehabilitation helps ebb regional recruitment and learned non-use for patients [109], the inclusiveness of BCI extends these benefits to children who could not use traditional PT. Early-life BCI intervention does not necessarily need to be used in isolation for therapeutic gains. In fact, supplemental use of BCI with physical therapy or other associated learning provides a strong option which may help reduce the physical demands on the patient while still promoting healthy rehabilitation [63, 74–76, 112]. Incorporating in the more approachable therapeutic option may help develop a more consistent and enjoyable rehabilitation scheme for the patient. While BCI technologies have their own trade-offs, the freedom from physical limitations allows the benefits of rehabilitation to be accessible to a wider scope of patients.

Beyond advantages in accessibility, BCIs provide an opportunity for more customizable motor rehabilitation schemes. The programmable nature of BCI allows a freedom to create and refine engaging applications to enhance neurorehabilitation approaches at functional and user-interface levels [108]. Functionally, customizing BCI for individual end users could include concepts like targeted rehabilitation and improved neuroplastic activation specificity, leading to more personal rehabilitation. At a user-interface level, presenting therapeutic activities under the guise of an engaging application or game is extremely

promising for development of BCI, especially for younger users [69, 118]. Increased user-interface accessibility to the rehabilitation paradigm can reduce user frustration and promote engagement, leading to improved patient rehabilitation [99, 108]. These advantages help justify translating BCI technology for early-life neurorehabilitation applications.

3.6.2 Prospective Solutions to Barriers Affecting BCI in Children

Synthesizing the literature evidence above, it becomes clear there is support for translating motor rehabilitation applications to children. Further, literature evidence presented throughout this chapter help outline potential barriers to translating motor rehabilitative BCI to children. Below, potential solutions to some of these barriers are explored alongside suggestions for research topics which could help develop BCI in children.

There are unique technical challenges facing EEG based BCI development for children. One technical challenge facing BCI in children is constructing a dynamic BCI system which can adapt to and handle the developing brain. In creating these rehabilitation BCI applications, understanding the underlying EEG characteristics of children is critical. EEG properties alter significantly throughout development, resulting in key EEG frequency bands, topography and power distribution shifting dramatically with age [21, 119, 120]. These alterations are more prominent for some types of BCI, e.g. MI-BCI, since they rely on parts of the EEG spectral landscape which change dramatically throughout childhood [30, 69]. Designing an EEG-BCI which can accurately interpret this flexible spectral landscape is a critical open question. The challenge is unique to child BCI research, but its solution may be applicable to improving other general BCI issues such as variance across sessions. This challenge could be approached through several different methods. For example, recent research developing adaptive boosting-algorithms [121] could be utilized to help appropriately identify MI signals from children. Considering the adaptive algorithm does not require a priori definitions for frequency bands [121], it is a potential option to partially circumvent issues like evolving frequency band ranges in the developing brain. Another option may be to incorporate higher dimensional information, like age, into the signal processing and analysis aspects of BCI through tensor techniques [41, 44, 122, 123]. Tensor methods may prove to be additionally useful to translating BCI to children as a means to remove artefacts [124], potentially mitigating some of the increased noise in children's EEG. Exploring topics like convolutional neural

networks [125] and deep learning structures [126] within the framework of the developing brain provide additional paths which could produce key developments useful in translating BCI to children. Building upon concepts like these for use in the developing brain, be it the flexibility in autonomously defining spatial-spectral configurations [121] or utilizing higher dimensional information in tensor methods [38], will likely be key. Research which focuses incorporating the dynamic EEG characteristics of children into machine learning and signal processing methods can help in resolving some of the technical aspects barring the translation of BCI applications to children.

Suitable development of engaging and accessible BCI applications is another challenge to focus on for BCI research for children. The benefits of this research could be applicable to older BCI users, but is critical for success for younger BCI users. Younger BCI users may have reduced attention spans and a potentially higher sensitivity to the fatigue associated with BCI. Long training sessions and mental strain in BCI are thus two areas of research which need to be addressed. A greater push for development of engaging age appropriate applications may partially resolve these concerns. Having an age appropriate paradigm embedded in the medium of play could influence sustained use by the user (like in [68, 113]) while less appropriate designs can lead to high task-completion failure rates (like in [115]). Engaging paradigms for the training component of a BCI may be especially important for neurorehabilitation applications, as improved training can greatly improve the accuracy of the BCI [99, 108].

Improvements in dry electrode hardware helps address the concern of attention in paediatric populations through reducing the initial set up and preparation time of EEG [127, 128]. Research providing methods to reduce the required number of electrodes needed for effective EEG based BCI [129] provides another opportunity which could be further explored to optimize BCI for children. These considerations are some examples of research areas which could be focused on to improve BCI applications for children.

Ethical considerations are also of particular concern when translating and designing BCI technology for children (for a general review of BCI ethics see [130], and human ergonomic considerations see [103,131]). This area is crucial to effective understanding of how to translate the BCI rehabilitation technology to children in a realistic manner. Concerns about long-term effects and any possible complications of sustained BCI use is currently unknown, and must

BCI literature featuring children				
Study	Number of patients (age range)	Pathology of interest	BCI methodology	Limitations
Lim *et al*, 2012 [113]	$N = 20$; (age 6-12)	ADHD	EEG (NFB)	Non-blinded study, potentially biased results
Rohani *et al*, 2014 [114]	$N = 5$; (age not disclosed)	ADHD	EEG (P300)	Ages not disclosed, paradigm needs redesign
Ehlers *et al*, 2012 [115]	$N = 51$;(age 6-33), $N = 37$; (age < 11)	Communication/BCI Speller	EEG (SSVEP)	Application design may have led to high failure rate for youngest subjects
Breshears *et al*, 2011 [116]	$N = 6$; (age 9-15)	Neuroprosthetic control	ECoG (MI)	Invasive, control derived from high gamma band
Roland *et al*, 2011 [117]	$N = 23$; (age 11-53), $N = 10$; (age < 25)	Neuroprosthetic control	ECoG (MI)	Invasive, control derived from high gamma band

Table 3.1: Part 1 of the summary table of key aspects from the literature review, highlighting literature which explicitly uses BCI with children.

be weighed heavily before implementation in a clinical setting. Also, it is imperative to clearly describe to both the children and parents alike each part of the BCI, such as the sensor array, its role and how the system will be run, to alleviate as many worries as possible prior to use.

These prospective solutions are not comprehensive, but present several significant challenges in technical execution, analytical development and realistic implementation when designing and developing BCI applications for the paediatric brain. A summary of these solutions and their prospective role in translation is provided in Tables 3.1, 3.2, 3.3. Fully realizing BCI applications for children is dependent on being careful and mindful of these challenges.

3.7 Update: A Return to EEG-BCI Literature for Children in 2018

Since our initial literature search on EEG-based BCI for children, an increasing amount of activity surrounding the development of paediatric BCI applications, at both engineering and clinical levels, has appeared. These research contributions span across the various BCI frameworks, but have largely focused on using EEG-based BCI. As some brief examples, both Qian *et al.* [132] and McLaren *et al.* [133] explore how to incorporate EEG-based BCI training for treating ADHD in children. The former describes introducing the BCI training through a gaming medium [132], while the latter uses sound and auditory treatments to help children self-regulate their attention [133]. Another group, led by Gurkok *et al.* have outlined and proposed a framework for developing BCI games [134], with the hopes to transfer knowledge from the game development and BCI communities into a shared framework for better collaboration and joint research. Norton *et al.* even references our original survey in their manuscript describing the performance of 9-11 year-old children using an SSVEP-based

Potentially applicable technical developments			
Study	Research Goal	Relevance to EEG-motor rehabilitation BCI	Relevance in translating motor rehabilitation in BCI to Children
Darvas *et al*, 2010 [14]	To demonstrate EEG can be used to acquire high gamma (HG) signatures through functional mapping of HG activity to a cortex model using MNLS and voxel-wise computed time-frequency maps for ROIs.	Provides access to task induced HG frequency signatures in EEG. Spatially localized HG power changes and interhemispheric phase synchronization signals derived from EEG using htis method were akin to ECoG values.	HG control in MI-BCI in paediatric populations using ECoG provides a stable signal robust to age [14, 116, 117]. Being able to access and analyze these signals using non-invasive EEG could provide a partial solution to dealing with the changing electro-physiological profile [22, 120] of the developing brain.
Liu *et al*, 2015 [121]	To autonomously select key channels and frequencies for stroke rehabilitation through an adaptive boosting algorithm applied to the usually pre-determined spatial-spectral configurations modelled as variable preconditions.	Provides details on a technique for training weak classifiers through a new heuristic supervisor of stochastic gradient boost strategy applied to preconditions leading to optimal spatial-spectral selection for BCI rehabilitation.	Autonomous spatial-spectral configuration can help address differences in the spectral landscape present throughout development [21,22,119,120]. EEG processing techniques which do not require a relatively broad or pre-determined frequency range and channel selection can also help indicate which spatial-spectral configurations are actively changing throughout recovery [121].
Liu *et al*, 2014 [44]	To detect MI-EEG patterns in spatial-spectral-temporal domains by a tensor-based scheme constructed using a wavelet transform method.	Tensor methods retain the multi-way (i.e. multi-mode) nature of EEG. Extraction by a tensor-based nearest feature line distance (TNFLD) algorithm and SVM classification allow greater separation of MI patterns from EEG in stroke patients.	Retaining the higher dimensionality in signal analysis provides options to incorporate developmental information, like age, to tackle differences in signal patterns and selection for BCI in children. Tensor factorization methods also provide a means to address the changing MI-EEG patterns present during rehabilitation therapy [44].
Zhang *et al*, 2016 [124]	To remove EEG artefacts for BCI through Bayesian tensor completion via specifying a sparsity-inducing hierarchical prior and automatically inferring model parameters of the underlying low-rank tensor through Bayesian inference.	Recovers the disturbed data in EEG recordings with artefacts, and uses possible outliers as missing values for EEG tensor completion. The artefact completion method provides additional information for BCI which is lost in typical artefact rejection.	Children are naturally prone to EEG artefacts and noise [22]. Retaining as much information as possible from EEG in children is important due to their shorter attention spans (and hence high artifact probability) and reduced signal power [119]. Converting EEG data with artifacts into usable information through signal recovery helps address these problems, leading to more information for BCI applications.
Lau *et al*, 2012 [129]	To determine how reducing the number of EEG channels affects electrocortical source signals that can be parsed from recorded EEG.	Provides a basis to examine how many EEG sensors are required for BCI applications through applying an adaptive mixture ICA algorithm on EEG channel subsets.	Reducing the required number of EEG channels can help reduce set-up time of BCI. This reduction may be critical for young BCI users due to their shorter attention spans. Consequently, more rapid set-up times could improve BCI interaction for children.

Table 3.2: Part 2 of the summary table of key aspects from the literature review, outlining potential technical developments for BCI in children.

Considerations on BCI user interaction and ethics		
Study	Research Goal	Relation to translating BCI to children
Friedrich *et al*, 2014 [68]	To design a combined BCI and biofeedback treatment for children with ASD based in play.	Provides an example for designing BCI applications specifically for extended rehabilitation in children, based around play. Obstacles in the development process help highlight the importance of clever paradigm design for implementing applications in children.
Nijboer *et al*, 2013 [130]	To discuss ethical issues related to BCI, its research and development.	Provides a starting place for ethical considerations which need to be identified for BCI applications in vulnerable groups, like children. Ethical questions are important to evaluate when designing rehabilitation applications, especially for young children.

Table 3.3: Part 3 of the summary table of key aspects from the literature review, emphasizing literature examples of considerations beyond technical improvements which could be critical to BCI application development. Acronyms for all parts of the summary table: Attention-deficit hyperactivity disorder (ADHD); Autism Spectrum Disorder (ASD); Electroencephalography (EEG); Electrocorticography (ECoG); High gamma (HG); Independent Component Analysis (ICA); Motor Imagery (MI); Minimum norm least squares (MNLS); Neurofeedback (NFB); Region of Interest (ROI); Steady-state visual evoked potential (SSVEP); Support Vector Machine (SVM); Tensor-based nearest feature line distance (TNFLD).

EEG-BCI [135]. The goal of Norton and associates investigation was to demonstrate a new approach to SSVEP-task selection using BCI in children, outside of the standard spelling task [135]. They highlight the potential downfalls outlined in this chapter, e.g. the failure rates seen in Ehlers *et al.* [115, 135], as evidence for the need to better design BCI tasks for children. Together, these examples broadly highlight some of the new publications in EEG-based BCI for children (since our original literature survey) which are not focused on motor rehabilitation tasks.

While our own contributions [2,3] comprise a decent portion of literature with paediatric motor rehabilitation as an end-goal motivation, other groups have shown complementary interests in this application in recent years. For example, researchers de Oliveira *et al.* [136] describe a novel virtual environment and playful therapeutic game for neuropsychomotor rehabilitation. Their work supports our initial comments on the importance of play to engage children in rehabilitation tasks, and highlights its usefulness for children with cerebral palsy. In their report, de Oliveira and crew incorporate EEG-based sensors to help monitor engagement and attention levels during the task execution, in the virtual space [136]. Their approach has been lauded by several clinical experts as a promising tool for rehabilitation in children with cerebral palsy [136]. Furthermore, at the most recent annual international Brain-Computer Interface Conference (BCI2018), there were five separate submissions focused on paediatric EEG-based

BCI applications, with several focusing on its use in motor rehabilitation scenarios [137]. These recent developments position our continued efforts in improving the ability to translate BCI motor rehabilitation applications to children as a valuable means of contribution to the field.

The growth in interest in BCI-based paediatric applications overall can be represented by the work of Moucek *et al.* [138]. In their report to *Scientific Data*, they offer a large, publicly available 'odd-ball paradigm' collection of P300-based BCI experiments and data collected on school age children [138]. Their work offers EEG recording, event stimulus and demographic information on approximately 250 (138 male, 112 female) school-age children and teenagers, between the ages of 7-17 ($mean = 12.9$) [138]. Their goal is to offer a large, easily accessible dataset which can be used as the basis for establishing comparisons between different studies focusing on this type of BCI work in children [138]. As such, this manuscript represents an ongoing, and concerted effort to help provide data for more researchers to develop BCI-focused applications for school-age children.

3.8 Limitations

The presented literature survey has several limitations. First, the major limitation of the review is the narrow definition used in the inclusion/exclusion steps for what is considered a BCI. In the original review, the literature searching explicitly required one of the terms 'BCI', 'BMI', 'brain-computer interface', 'brain-machine interface', 'neurofeedback', etc. to be included. While the initial inclusion step tried to be lenient, the narrowness of needing one of these specific terms necessarily restricted the types of papers found. It is possible that experiments in other relevant fields, such as Psychology, may be reporting on brain-computer interfaces for children, but these were not found in the literature survey due to a difference in how different research fields term such experiments. As such, there are potentially additional papers not covered by this review which explore similar topics, but from a different field. This effect was slightly mitigated during the second part of the literature search, in which relevant citations of papers included from the initial search were evaluated as part of the analysis. However, it is important to note that the restrictive nature of looking for these terms may have limited the relevant manuscripts to only more engineering related fields. Ideally, in future work unique nomenclature for similar topics across various fields could be incorporated into such

a structured review as to not exclude relevant manuscripts.

Similarly, in the original review, virtual reality was not explicitly included in the analysis. This choice was made when the original literature survey was conducted due to the potentially mixed levels of subject engagement, e.g. simply exploring a virtual space vs. actual brain-computer interaction. However, with the increasing accessibility of the virtual and augmented reality (VR/AR) systems, and significant advancements in synchronizing the EEG-based BCI signals to these digital platforms, VR/AR applications in BCI have gained significant popularity in recent years. As such, they should be evaluated more closely when identifying literature evidence supporting BCI applications in children going forward.

The review is also limited when drawing some conclusions based on post-stroke studies for evidence while other motor impaired conditions, like cerebral palsy, are more common in children [78, 79]. As motor rehabilitation BCI becomes more accessible and is applied across different pathologies, a re-evaluation of this restriction can be more appropriately examined. This review is also limited when drawing conclusions about appropriate age ranges of subjects for BCI due to the limited literature and resources available. Additional studies are needed in order to more definitively designate specific age ranges that could use different types of BCI technology.

3.9 Conclusions

Therapeutic BCI applications have shown great strides in neurorehabilitation and the recovery of motor deficits for patients. These advances in rehabilitative technology have yet to manifest in therapeutic BCI applications for children. The original survey [1] examining literature evidence clearly illustrated a positive trend which suggested that developing motor rehabilitation applications akin to those seen in adults could be beneficial for children [68, 77, 82, 98, 109, 116]. Through exploring the results of associated literature recovered from a systematic search, insights into concerns and barriers affiliated with bringing motor rehabilitative BCI to children were discussed. Literature from ECoG applications with children highlighted that functional control of a BCI was definitively possible for young users. Prospective solutions to some of the main methodological barriers in EEG-based BCI were suggested, including the technical foundation for the remaining chapters in this thesis, with

recommendations given for avenues of potential future work. Revisiting the current literature three years after the original survey finds increasing interest on incorporating child-focused designs and applications for EEG-based BCI paradigms.

Therefore, the presented literature evidence supports and motivates this thesis as timely and significant with respect to the goal of helping translate EEG-based BCI applications to children. However, in order to fully realize the BCI applications in children, the associated challenges of BCI in children need to be addressed. The remainder of this thesis aims to address a set of these barriers from a signal processing viewpoint. In this way, the contributions of this thesis would fit alongside the work by Liu *et al.* [44] in Table 3.2, extending the tensor framework to paediatric EEG signals. Through establishing a multi-way framework to better identify and characterize common and latent structures present in paediatric EEG signals, we provide a set of tools which offers solutions to the barrier of accounting for the heterogeneous, dynamic child brain development.

Chapter 4

Characterizing Developmental Feature Profiles Using Multi-way Analysis in Paediatric Resting-state EEG (rEEG)

This chapter is built upon the published submission to the 2017 *IEEE Engineering in Medicine and Biology Conference* (EMBC-2017) [2], and the subsequent elaboration of the work into a published journal article in the *Journal of Neural Engineering* in 2018 [3].

The key contributions in this chapter include:

- Developing a technique to identify key number of components and relevant features based on unimodality constraints and CORCONDIA
- Using multi-way analysis on rEEG to capture development-specific features across ages
- Demonstrating the multi-way technique's ability to predict age better than random chance or naive classification.
- Validation on 4 separate paediatric datasets and a simulated dataset.

4.1 Introduction

EEG signals from young children can be confounded by development-related phenomena, as discussed in Chapters 2 and 3 [12, 22, 28]. Differences in brain maturation can lead to varying EEG rhythms, high background noise and obfuscated signals of interest [1, 12, 21, 22, 28]. Such differences can be partially dealt with *a posteriori* for each child, using techniques like isolating individual alpha frequencies to identify the location of specific EEG rhythms of interest [139],

but these techniques do not alleviate other issues facing paediatric EEG analysis. Therefore a data-driven means to detect and extract latent EEG features related to development which mitigate some of these challenges in EEG analysis could be a useful tool.

This chapter provides a proof-of-concept for characterizing developmental feature profiles of children using multi-way analysis on paediatric rEEG data. By exploiting the multi-way connections inherent in paediatric EEG, tensors offer a potential framework to capture the underlying developmental links common throughout childhood and across groups of children. Here, a robust feature selection paradigm is presented, which includes tensor component selection and model validation. The developed paradigm is demonstrated on several resting-state paediatric EEG datasets which span:

- a early-life pre-school population with potential developmental impairments.
- a population which spans childhood to adulthood with potential developmental impairments.
- a healthy population during a stable developmental period of childhood.

Successful characterization of latent age-specific features in each resting-state dataset supports the proposed framework as an adaptive tool for incorporating development-sensitive feature selection into paediatric EEG signal processing.

4.2 Materials

4.2.1 Datasets

Muir Maxwell Epilepsy Centre

A retrospective analysis of an epileptic preschool cohort (< 5 years) from the Muir Maxwell Epilepsy Centre (Edinburgh, UK[1]) was included in this study, henceforth referred to as the MMEC dataset. The original cohort was prospectively recruited from National Health Service (NHS) hospitals in Fife and Lothian as part of the NEUROPROFILES study [140]. A 32-channel, unipolar montage [10] captured routine EEG in the standard 10-20 system for

[1]Thank you to Matthew Hunter, Michael Yoong and Richard Chin of the Muir Maxwell Epilepsy Centre for providing access to this data and granting permission for its use in our analysis.

each child at a sampling rate of 511 Hz. EEGs were recorded at 20 scalp electrodes (FP1, FP2, FPz, F3, F4, F7, F8, Fz, C3, C4, Cz, P3, P4, Pz, T3, T4, T5, T6, 01, 02), eight auxiliary electrodes (AUX1-8), two grounding (A1, A2) and two ocular electrodes(PG1, PG2). Of 64 children available, 14 were excluded from this study due to corrupted EEG data, inconsistent or incompatible EEG acquisition parameters and irregular recordings, resulting in a dataset of routine EEG from $n = 50$ preschool children. If multiple resting-state EEG recordings existed, only the first recording was selected for each child. This avoided weighting results toward children with more recordings. Data was selected from similar awake resting-state recordings across all children.

Children's Hospital Boston-MIT

Publicly available data from the Children's Hospital Boston-MIT of epileptic patients from infancy to early adulthood [141] was used in this study, downloaded through Physiobank.org [142] and henceforth referred to as the CHB-MIT dataset. A 28-channel, bipolar montage captured EEG recordings continuously over two days of monitoring at a sampling rate of 256 Hz. Of 23 available patients, montage discrepancies and unsuitable recordings in 6 subjects were deemed to be too inappropriate for inclusion in this study. This resulted in $n = 17$ subjects (age 2-19 y.o.) for analysis. The 48-hour continuous recordings were separated into 4-hour time blocks, with an equal number of trials at each time selected for processing. Results were averaged across all time bins to render a holistic representation of the resting-state data for each subject.

Child Mind Institute

Resting-state EEG data for healthy control participants was taken from the open science resource provided by [25] and the Child Mind Institute. EEG data was recorded from a high-density 129-electrode hydro-cel EEG cap at a sampling rate of 500 Hz, and filtered by the CMI at a bandpass of [0.1 - 100] Hz [25]. Data captured from the resting-state paradigms of pre-adolescent subjects (one age 6, the rest age 8-11 y.o.) was used. Of 45 subjects available, one subject was excluded (age 11) due to the EEG being recorded using a different montage, resulting in $n = 44$ subjects for analysis. The single 6-year-old in the dataset was grouped into the 8-year-old class to allow for cross-validated classification (see Section 4.3.6).

A summary of subject distribution per age for each dataset is included in Figure 4.1.

4.3 Methods

4.3.1 Pre-processing

Raw EEG data was processed using the Fieldtrip toolbox described in Section 2.8.1 [58]. A two-pass (zero-phase forward and reverse) bandpass filter between [0.5-31] Hz was applied to recorded EEG data. The signals were separated into 10-second (2-second) long trials for the MMEC, and CHB-MIT (CMI) datasets. Both the 10-second and 2-second trials were recorded resting-state EEG, as described in Section 2.3.3. As such, the short recording windows used for analysis were capable of reflecting the non-specific brain processes involved with the resting-state condition, with no significant differences due to the distinct trial lengths. The recordings were detrended in order to remove linear trends which could affect results in the low frequencies. Trials in the CMI were selected to be shorter in order to improve computational speed of the analysis.

EEG channels were re-referenced to a common average reference and auxiliary/reference specific channels were removed. The EEG channel montages in the MMEC and CHB-MIT were both recorded with respect to reference channels. The CHB-MIT montage used a so-called 'banana' EEG montage, recording the electric potential as a joint combination of two EEG electrodes. This pairwise electrode montage from the CHB-MIT was adapted to the unipolar montage of the MMEC, where EEG potential was described by single electrodes instead of pairs. The high-density set-up of the CMI data was not adapted to this same convention to avoid potential information loss. Channels in the CMI data without measurable numeric values, e.g. data containing Not-a-Number (NAN) values produced by masking artefacts in the pre-processing stage, were removed. Trials with seizure activity were documented by clinicians prior to us receiving data. These trials were not included for processing.

A multi-pass artefact rejection system was used. First EEG electrode artefacts such as those produced by muscle movements, electrode shifts (i.e. electrode 'jump' artefacts) and ocular blinking artefacts were automatically identified using strict rejection criteria relative to the

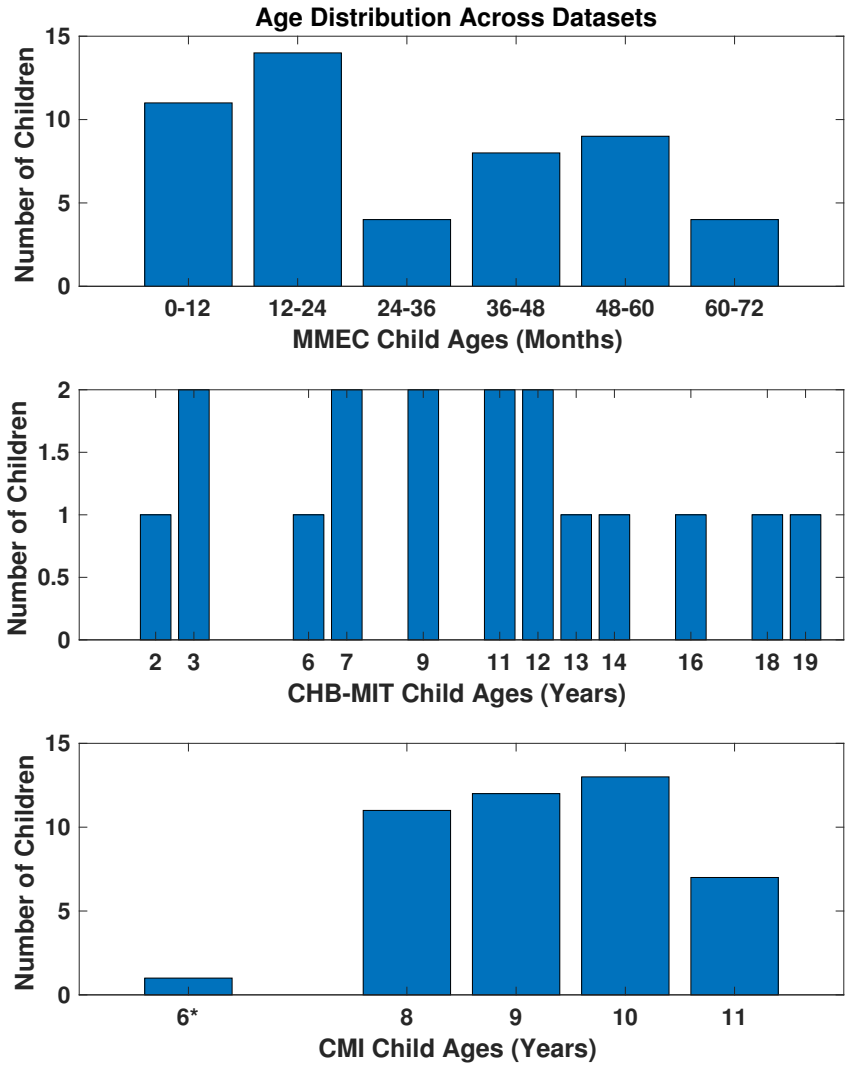

*Figure 4.1: Distribution of subjects per age for the MMEC, CHB-MIT and CMI datasets. *The single six-year-old child in the CMI dataset was included in the eight-year-old category for classification.*

Fieldtrip default suggested values [58] (Fieldtrip release range R2015-R2016b, z-value rejection level $r = 0.4$). Then manual inspection of the data was used to identify and remove any additional artefacts potentially missed by the automatic rejection. Manual artefact rejection removed outliers in both trial and channel data based upon high variance values ($var > 10^6$).

4.3.2 Tensor Construction

Three-way tensors consisting of $[Spatial] \times [Spectral] \times [Subject]$ modes were created using the EEG channel, power spectra and subject age data for each dataset. This resulted in three dataset tensors with elements corresponding to the number of channels, number of samples in the power spectra and number of subjects. The MMEC had 19 channels by 301 samples in the power spectra by 50 subjects, e.g. $(19) \times (301) \times (50)$ elements in the tensor. The CHB-MIT tensor consisted of 19 channels by 301 samples in the power spectra by 17 subjects, e.g. a tensor with $(19) \times (301) \times (17)$ elements. The CMI tensor was comprised of 105 channels by 61 power spectra samples by 44 subjects, e.g. $(105) \times (60) \times (44)$ elements in the tensor. Figure 4.2 (A) provides a general illustration of the 3-way tensor construction.

Time-frequency analysis was done on artefact-free trials using a filtering window method from the Fieldtrip toolbox presented in Section 2.8.1. The analysis utilized the multitaper method from Fieldtrip with 0.5s Hanning window. Normalized power spectra were recorded for each subject in the MMEC and CHB-MIT datasets at 0.1 Hz resolution. Power spectra were calculated at 0.5 Hz resolution for the CMI dataset as a result from the shorter trials in pre-processing the data. A trial length of 10 seconds was explored for the CMI dataset, to examine if it would affect the rEEG analysis. The shorter trial length of 2 seconds and its altered frequency resolution was found to have no appreciable difference for the proposed analysis, so the shorter window length was kept. The power spectra in each subject were averaged across all trials to provide a general spectral profile of the resting-state EEG for the $[Spectral]$ domain.

Children in the $[Subject]$ domain were ordered in each dataset to be increasing in age based on the available data. For the MMEC dataset, each child's age was given as months old so children were ordered by age from youngest to oldest in months. For the CHB-MIT and CMI dataset, children's age was given in years. Thus, they were ordered from youngest to oldest

by years, with a consistent ordering used in repeated analysis. The structure of the $[Subject]$ domain can then function as a proxy for child development. Through this, age-specific features can be readily identified based on their associated position within the $[Subject]$ domain.

4.3.3 Tensor Factorization: Constraints and the CP Model

The CP model described in Chapter 2 was used as the basis for the tensor factorization of each dataset. Figure 4.2 (B) illustrates a general pictorial representation of a tensor decomposition for varying components, R, for a 3-dimensional tensor.

The CP model was chosen here in order to exploit its guaranteed 1:1 interaction between extracted factors across the given modes, as described in Section 2.5.3. Due to the CPs restricted interaction between only corresponding factors, we can gain insight into the structural relationships within the constructed tensors, in particular how increasing age in the $[Subject]$ domain is related to other modes. For example, in Figure 4.2 (B) we can gauge how the first factor (*red* in $R(1)$) in the $[Spatial]$ domain corresponds directly with only the first factors in the $[Subject]$ (*green* in $R(1)$) and $[Spectral]$ (*blue* $R(1)$) domain. This guaranteed interaction from the CP model is critical when evaluating the latent developmental relationships present throughout childhood, and verifying their existence across different sets of children. Furthermore, being able to retain uniqueness using the CP model under only mild requirements (Section 2.5.3) permits easier interpretation of how the developmental $[Subject]$ domain influences factors in the $[Spatial]$ and $[Spectral]$ modes. As such, the CP model provides a foundational support to the investigations in this chapter.

Constraints on the modes of the CP model were used to improve interpretation of results, analogous to the use of constraints in the NNMF [33, 34]. Constraints also helped retain mode-specific properties [33], e.g. non-negativity in the power spectra. Non-negativity was enforced in the $[Spatial]$ and $[Spectral]$ domains. Enforcing non-negativity in these domains ease the amount of interpretation needed in evaluating how specific elements are participating in the CP model decomposition. As mentioned in Section 2.5.6 evaluating the outputs from the CP model requires interpretation of the relationship between the data basis of the original tensor. By including non-negativity for a domain, such as the $[Spatial]$ domain, the interpretation is simplified to a value-added model. For example, consider the non-negativity

imposed on the $[Spatial]$ domain. With only positive values considered in the resulting factor matrix of the $[Spatial]]$ domain we can interpret how much value was added to the model from each EEG electrode. Put another way, the non-negativity allows us to easily gain insight into how important specific channels are for reconstructing the data based on how much value was added by each electrode (e.g. 0 indicates almost no contribution to the model). Negative values in the $[Spatial]$ domain, however, would require more complex interpretations where it is necessary to determine what negative values indicate at a given electrode. Considering the CP model is not reflecting the raw electrophysiological potential, but rather the electrodes' relationship to the $[Spectral]$ and $[Subject]$ domains, the concept of negative values in the $[Spatial]$ domain become more complex. As such, to ease of interpretation of outputs, this thesis uses non-negativity as a common constraint on most modes throughout the proposed analysis.

Unimodality was implemented in the $[Subject]$ domain, via the built in function in the N-way toolbox for Matlab [59]. The unimodality constraint follows the unimodal least squares regression (ULSR) algorithm from [143], where the ULSR is non-negative as well as unimodal. In brief, the ULSR solves the problem:

$$\min_{\mathbf{a}} ||\boldsymbol{\alpha} - \mathbf{a}||_F^2 \tag{4.1}$$

subject to **a** as unimodal. Solving this problem proceeds as follows [143]:

- First, suppose a monotonic increasing regression is performed on $\boldsymbol{\alpha}$. Let, $\mathbf{a}^l$ be the monotonic increasing regression on $\boldsymbol{\alpha}$, and $\mathbf{a}^D$ be the monotonic increasing regression on $\boldsymbol{\alpha}$'s reversed order elemental order.
- Then, let $\mathbf{a}^{l,n}$ be the monotonic increasing regression on the first $n-1$ elements of $\boldsymbol{\alpha}$, and $\mathbf{a}^{D,n}$ be the monotonic decreasing regression on the last $J-n$ elements of $\boldsymbol{\alpha}$ [143].
- Next, for all n, from $n = 1, \ldots, J$ define:

$$\mathbf{c}^{(n)} \equiv \begin{bmatrix} \mathbf{a}^{l,n} \\ \boldsymbol{\alpha}_n \\ \mathbf{a}^{D,n} \end{bmatrix} \tag{4.2}$$

- Then, solving for **a** by:

$$\mathbf{a} = \underset{\mathbf{c}^{(n)}}{\operatorname{argmin}} \left(||\boldsymbol{\alpha} - \mathbf{c}^{(n)}||_F^2 \quad | \quad \max(\mathbf{c}^{(n)}) = \boldsymbol{\alpha}_n \right) \tag{4.3}$$

  provides a means to select a $\mathbf{c}^{(n)}$ which minimizes $||\boldsymbol{\alpha} - \mathbf{c}^{(n)}||_F^2$ by selecting only the $\mathbf{c}^{(n)}$ satisfying $\boldsymbol{\alpha}_n = \max(\mathbf{c}^{(n)})$ [143].

- Non-negativity is then enforced by setting all negative values in the regression vector to zero, based on the proof of Kruskal's monotonic regression and subsequent Lemma 1 [143].

These steps outline the solution to the ULSR algorithm in the N-way toolbox [59], thereby providing a unimodality constraint for the CP tensor data [143]. See [143] for more detailed description of the unimodality constraint, and accompanying figures.

Using unimodality was critical in order to guarantee that the extracted components were bound to only specific age groups within a dataset, and to help identify 'common ground' brain activation across subjects during resting-state. The $[Subject]$ domain structure was well suited for a unimodal constraint, as subjects had no repeated ages at the scale of 'months-old' and were arranged from youngest to oldest. Data was only grouped into 'years-old' sets for classification.

Tensor datasets were factored using our adapted PARAFAC function from the N-way toolbox. The adapted PARAFAC function extended the N-way toolbox script, improving its functionality when using both non-negativity and unimodal constraints in decompositions. As the N-way toolbox provided straightforward access to the unimodal constraint option, it was selected for our analysis.

4.3.4 Tensor Factorization: Component Selection

In the present work, an unknown number of underlying developmental relationships may be present in the data. Additionally, little *a priori* information is available to determine the best number of components which could capture the underlying structural relationships present throughout child development. The aim in this chapter is to account for the maximum number of potential developmental structures present in the data while retaining acceptable model

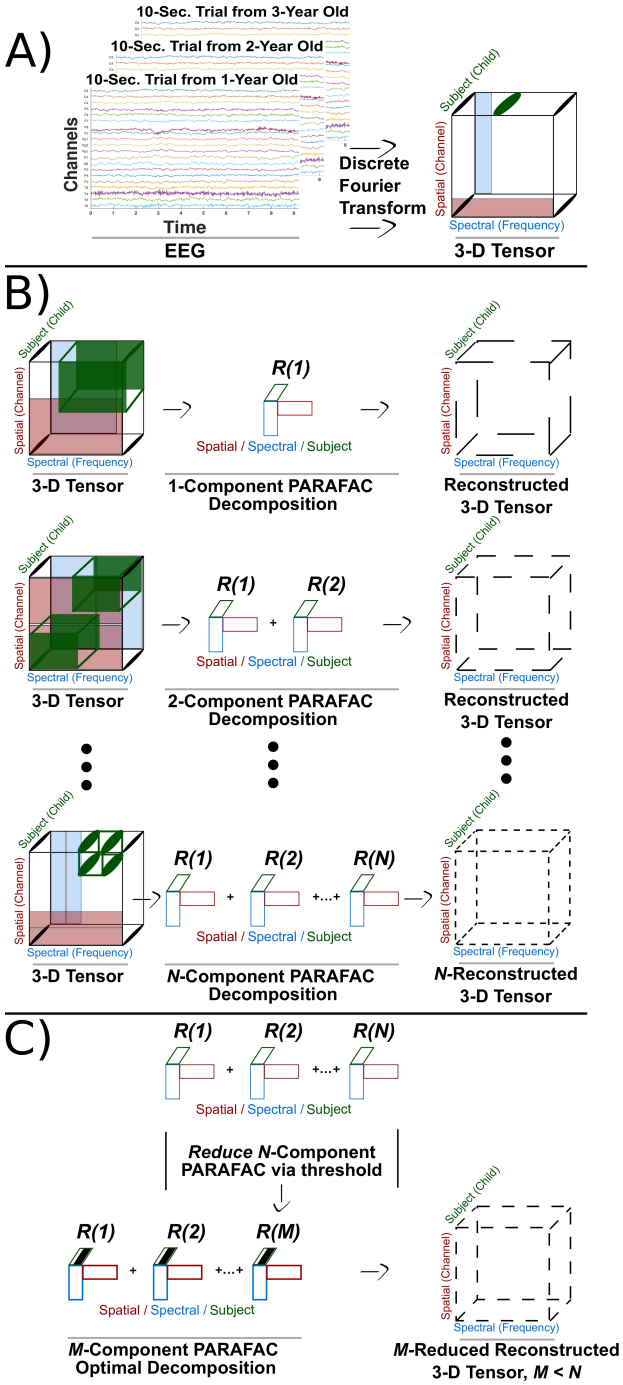

Figure 4.2: CP tensor construction and decomposition flowchart. A) Constructing a 3-dimensional tensor from raw resting-state time-series EEG with $[Spatial] \times [Spectral] \times [Subject]$ *domains on the* X, Y, Z *axes respectively. The* $[Subject]$ *domain is ordered to increase strictly with age. B) Illustration of* $R = 1{:}\,N$ *component grid search, where each* R*-factorized model is used to reconstruct the original data to test CORCONDIA and model suitability. C) Reduction of* R_N*-components to a optimal subset* R_M*-components (with* $M < N$*) based on a thresholding of the* $[Subject]$ *domain and model validation via CORCONDIA.*

viability. This can be achieved through exploiting the various properties of CORCONDIA, introduced in Section 2.5.5.

In particular, CORCONDIA can provide guidance towards identifying when the underlying relationships in a tensor are represented at a nearly 1:1 ratio with the number of components chosen for the CP model [48]. Consider again the tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ with $\mathbf{A} \in \mathbb{R}^{I \times R}, \mathbf{B} \in \mathbb{R}^{J \times R}, \mathbf{C} \in \mathbb{R}^{K \times R}$ as factor matrices. As the number of components $R = 1, 2, ...$ vary for a given CP decomposition, recall from Section 2.5.5 that CORCONDIA will also vary, as it naturally reflects how closely an estimated core tensor $\mathcal{G}$ is to the ideal super-diagonal core of a CP model. Then, for some given $r \in R = 1, 2, ...$ there will be a near 1:1 matching between the actual underlying structural relationships in $\mathcal{X}$ and the number of components selected [48]. At this point, all potential structural relationships in the data can be considered as 'fully represented' by the component factors, where there is a unique component factor to describe each unique interaction in the data structure. This state corresponds to the maximum number of component factors needed to uniquely capture all of the structural relationships, in other words the maximum suitable low-rank decomposition.

Thus, CORCONDIA at this fully represented point would find a core tensor $\mathcal{G}$ with significantly less interactions occurring off its super-diagonal, which would more closely represent an ideal CP core tensor and would lead to CORCONDIA naturally increasing [48]. Conversely, for all components $r + 1, r + 2, ...$ greater than r, the core tensor $\mathcal{G}$ would necessarily reflect more interactions off its super-diagonal core, as unique interactions in the underlying structural information would be shared by two or more component factors in $\mathcal{G}$. This in turn guarantees that the CORCONDIA at some ideal r will strictly be greater than CORCONDIA for any $r + 1, r + 2, ...$ etc. This result from CORCONDIA provides us with a critically important tool, as it gives us a threshold which can be used to identify the maximum number of unique components required to capture all latent relationships in the data uniquely. This solves our concern of having an unknown number of underlying developmental relationships and an unknown number of components needed to best capture these relationships in the resting-state EEG data.

In practice, a range of components $R = 1, 2, ..., 20$ was first used for the CP decomposition of each dataset, with CORCONDIA and explained variance recorded for each R. The maximum

rank decomposition $r = R$ (as outlined above) was chosen to be the maximum R which had CORCONDIA reported at 70% or greater. Then, a thresholding method was applied to exploit the CORCONDIA property described above, in order to determine the R-factor decomposition which would fully represent the latent developmental interactions in the data.

Thresholding was done using sample excess kurtosis values of the components in the unimodal $[Subject]$ domain. Kurtosis was calculated for $[Subject]$ mode factors with normalized loading values recorded above 0.6. Factors above 0.6 were determined to likely be influenced by a single child rather than a set of children. This was concluded from decomposition results using a large ($R > 5$) number of components. Kurtosis tested if the span of data influencing a factor in the $[Subject]$ mode was from a single child. Factors with values above 15 for the sample kurtosis were considered to be reflecting only a single child, and consequently excluded. Factors below the cut-off threshold were considered representative for at least a small set of children, and were included in the analysis. Through this thresholding method, only the subset of factors which included significant contributions spanning several children in the $[Subject]$ domain were selected. The reduced CP model then had its CORCONDIA re-evaluated. The reduced model with the largest number of components, R, which maintained CORCONDIA above 70% was then selected for analysis and training.

After thresholding, some sharp peaks may remain in the $[Subject]$ domain due to the unimodal constraint. These peaks correspond to the child who contributed most to the unimodal component factor, but do not imply the extracted factor only represent features from that single child.

Sets of 5 CP decompositions were run repeatedly under the same conditions (e.g. number of components) to help account for potential CP convergence to local minima. Figure 4.2(b) and Figure 4.2(c) illustrate the varying component $R = 1, 2, ..., N$ search, the CORCONDIA evaluation and the model selection process. Figure 4.3 shows full and reduced tensor model viability based on CORCONDIA calculations for components $R = 1, 2, ..., 19$ for the cross-validation folds of the MMEC dataset.

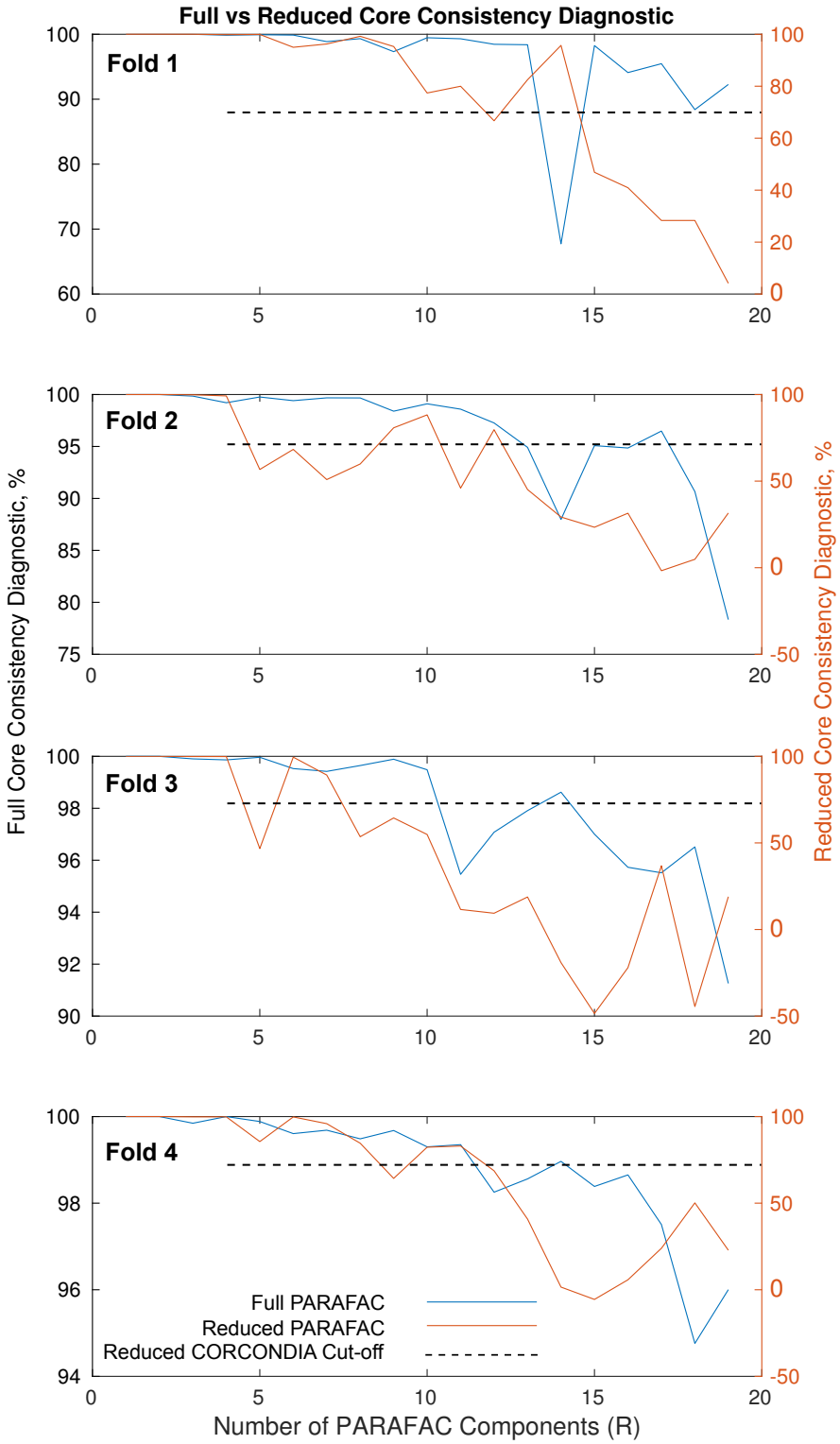

Figure 4.3: Full and threshold reduced CP model CORCONDIA for each fold in a factorization of MMEC dataset. Only a decomposition maintaining 70% CORCONDIA for both models was be considered viable, with the reduced tensor used for classification.

4.3.5 Classification: Modified Direct Projection

In order to evaluate the CP model's ability to identify and extract relevant developmental factors from the various resting-state EEG data, we employed the use of characteristic filters and direct projection, as described in Section 2.6. The advantages provided by the characteristic filters offered us a means to rapidly assess data in the testing cross-validation folds when evaluating classification in the MMEC and CMI datasets.

Recall in our given general direct projection equation (2.23) there is a pseudo-inverse step to estimating the characteristic filters. The pseudo-inverse introduces negative values in the newly estimated component weights. Given that in the present chapter, the CP model includes a strict non-negative constraint on the decomposed training tensor, the estimated component weights must also be non-negative in order to have meaningful predictive values for the new subjects. Therefore, the non-negative least square (NNLS) solution of the Khatri-Rao product in the encoding matrix $\mathbf{W}$ is introduced as an alternative to the pseudo-inverse step in (2.23):

$$\hat{\mathbf{C}} = \mathbf{Y}^{(3)} NNLS(\mathbf{W}), \tag{4.4}$$

with the NNLS function defined from Matlab as

$$\min_{x} ||W \cdot x - y||_2^2 \quad \texttt{subject to} \quad x \geq 0 \tag{4.5}$$

By using the NNLS solution, the non-negative constraints on the $[Spatial]$ and $[Spectral]$ domains are maintained. Values calculated for the estimated weights of $\hat{\mathbf{C}}$ in equation (4.4) are the basis for our classification analysis in this chapter.

4.3.6 Classification: Cost Matrices and Support Vector Machines

To maintain stringent integrity for classification, data was split into training and testing cross-validation folds prior to tensor decomposition. Therefore no testing data was included in building the training tensor. The number of folds for cross-validation were selected to retain stratification of the data in every training/testing combination, e.g. each fold has at least one representative from each class. The distribution of subjects per age in the MMEC and CMI

datasets therefore allowed for 4-fold and 5-fold stratified cross-validation, respectively. The CMI classification included the single six-year-old as a member of the 'Age 8' class in order to keep a stratified cross-validation. Therefore, for the $n = 50$ children in the MMEC dataset, with a 4-fold cross-validation, there were approximately two folds with 38:12 and two folds with 37:13 ratios representing the number of available training:testing subjects. Similarly, for the $n = 44$ children in the CMI dataset with a 5-fold cross-validation, there was one fold with 36:8 and four folds with 35:9 ratios of available training:testing subjects. Comparative classification within the CHB-MIT data was not possible due to the limited number of subjects per age.

A multi-class, ordinal classification scheme was devised to evaluate if the estimated component weights could predict subject age using the WEKA toolbox [61, 144]. Subject age (in years) was used for within dataset class labels. An ordinal cost-matrix was used to account for the multi-class, ordered nature of the data tensors. Misclassification penalties of the cost-matrix were linearly weighted, based on differences in class age. This had the effect of increasing classification penalties for predicting subjects as drastically older or younger compared to their actual age. Table 4.1 shows an example ordinal cost-matrix for classification between five age categories. The size of the cost-matrix was matched to the number of unique subject ages present in each dataset (see Figure 4.1).

A support vector machine (SVM) using a non-linear radial basis function (RBF) kernel and the ordinal cost-matrix was used to train a classifier from each fold on the CP decomposition factor matrices. The RBF-based SVM was selected due to the relationship between the extracted factors and child age appearing to be non-linear. Two hyper-parameters for the RBF-SVM were optimized using a grid search paradigm in WEKA [61, 144]. The parameters optimized were the complexity (C), and the influence rate of each data point on the support vector (γ). In a nutshell, the complexity hyper-parameter (C) determines the class margins and separation boundary definitions in the SVM by controlling how many data points are used as support vectors [61, 144, 145]. The influence rate hyper-parameter (γ) helps determines how influential single training examples are in the support vectors [61, 144, 145]. The WEKA toolbox then assessed varying combinations of C and γ by a grid-search implemented via nested cross-validation to determine the hyper-parameter values. Results reported here are

Ordinal Classification Cost Matrix

		Predicted Age Class				
		Age 0-1	Age 1-2	Age 2-3	Age 3-4	Age 4-5
True Age Class	Age 0-1	0	1	2	3	4
	Age 1-2	1	0	1	2	3
	Age 2-3	2	1	0	1	2
	Age 3-4	3	2	1	0	1
	Age 4-5	4	3	2	1	0

Table 4.1: An example of an ordinal cost-matrix for five age categories of children, from 0-5 years old. Penalty costs (weights) for misclassified ages increase linearly, moving away from the true class age.

evaluated on their overall classification accuracy and total penalty costs (e.g. the sum of all misclassification penalties based on the ordinal cost-matrix outlined in Table 4.1).

Two additional classification schemes were included for comparison to the RBF-SVM method. First, Random Selection (RS) classification was included, where classes were randomly assigned to subjects for each dataset. Additionally, a 'Single Class' (SC) classification (also known as a naive classification) was included, where only one class was selected for all subjects. Overall accuracy and total penalty costs were calculated for both the Random Selection and Single Class classification schemes. Results from each classification scheme are reported as averages with standard deviation across all training folds. A one-way ANOVA, corrected post-hoc with Tukey's Honest Significant Difference (HSD) for multiple comparisons, was used to infer differences between the RBF-SVM classification compared to the RS and SC classification schemes on the estimated projection data. A two-tailed Student's *t*-test was used to infer differences from the RBF-SVM penalty-costs as compared to the SC classification penalty-cost.

4.3.7 Visualization

To complement classification, results from the factorized training folds for the MMEC and CMI datasets are displayed using t-distributed Stochastic Neighbour Embedding (t-SNE) [146]. Using t-SNE, high-dimensional data can be visualized to capture both the local and global structure of the data through the presence of clusters. Demonstrating t-SNE maps on individual training folds in the data thus offers a visual companion to the classification analysis. The t-SNE maps show the local and global structure underlying a single fold present in the analysis.

4.3.8 Simulations

A simulation of pseudo-EEG data accompanies the real-world datasets. The Berlin Brain Connectivity Benchmark (BBCB) simulation code [147] was modified to include a shifting spectral frequency band of interest, similar to the developing alpha present throughout childhood [12, 21, 22, 28]. Lower bounds to the spectral signal band of interest were set as the mean simulated age $\pm$ 1 standard deviation. This was done for simulated subjects up to 8-years old. Afterwards, each additional simulated year was used to increase the sampling mean. This had the effect of gating the lower bound towards 8-Hz. The upper bound was set at 3-Hz $\pm$ 2 standard deviations higher than the lower bound. Variation in the upper bound was corrected to be at least 1-Hz above the lower bound. Ten children were simulated per age using the modified code, from 5 to 11 years-old. Simulated EEG was converted to Fieldtrip for processing, with a simulated tensor constructed in an identical fashion to the real-world datasets.

4.4 Results

The proposed tensor analysis successfully identified latent developmental features across subjects independently for each dataset. A detailed visual breakdown of the CP model with the resulting 'developmental feature profiles' is given in Figure 4.4 using the MMEC dataset. Qualitative developmental feature profiles are illustrated in each tensor domain for visualization purposes. The actual component weight values in the testing folds were used in the classification. Individual factor contributions are shown in the extended profiles to help clarify the latent developmental relationships present in each mode. Component profiles in the $[Subject]$ domain reflect in which ages the extracted factor (feature) is most dominant and influential. Features of the $[Subject]$ domain are ordered to match the highest to lowest explained variance from the $[Spatial]$ domain. The normalized topographic map of the $[Spatial]$ domain shows relative regional contributions of EEG channels for each developmental feature 'profile'. The $[Spectral]$ domain is shown up to 15 Hz, as higher frequencies for preschool children in resting-state data has little activity of interest and remains fairly flat.

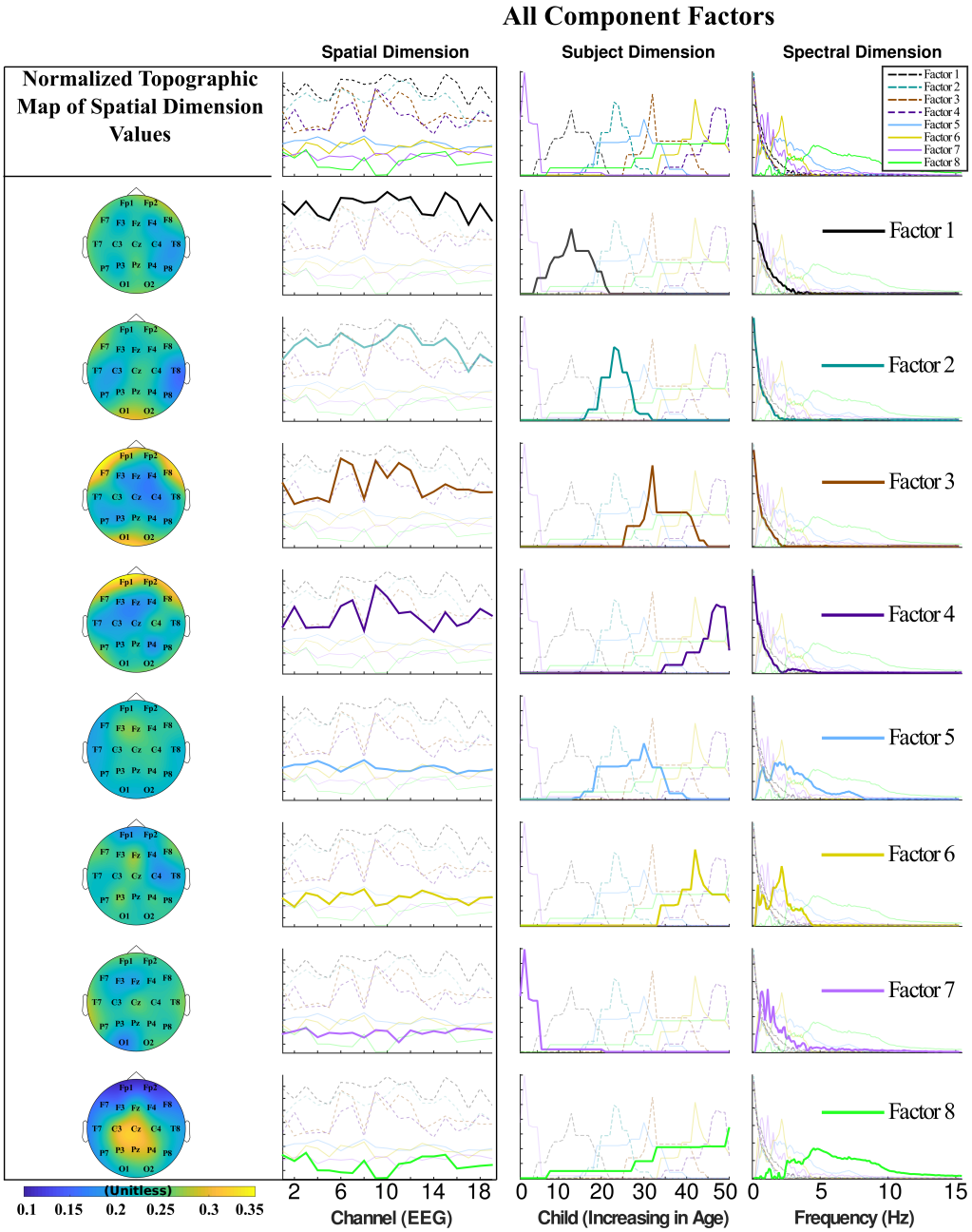

Figure 4.4: Detailed visualization of CP model extracted developmental feature profiles of epileptic children from 0-5 years old from the MMEC dataset. Normalized topographic maps of the $[Spatial]$ domain in column 1 show EEG channel regions with higher/lower relative contribution for each individual feature. The $[Subject]$ domain x-axis is the child's number (e.g. Child 1, Child 2), organized by increasing age. The combined and separated feature profiles for the $[Spatial] \times [Subject] \times [Spectral]$ domains are shown in columns 2-4.

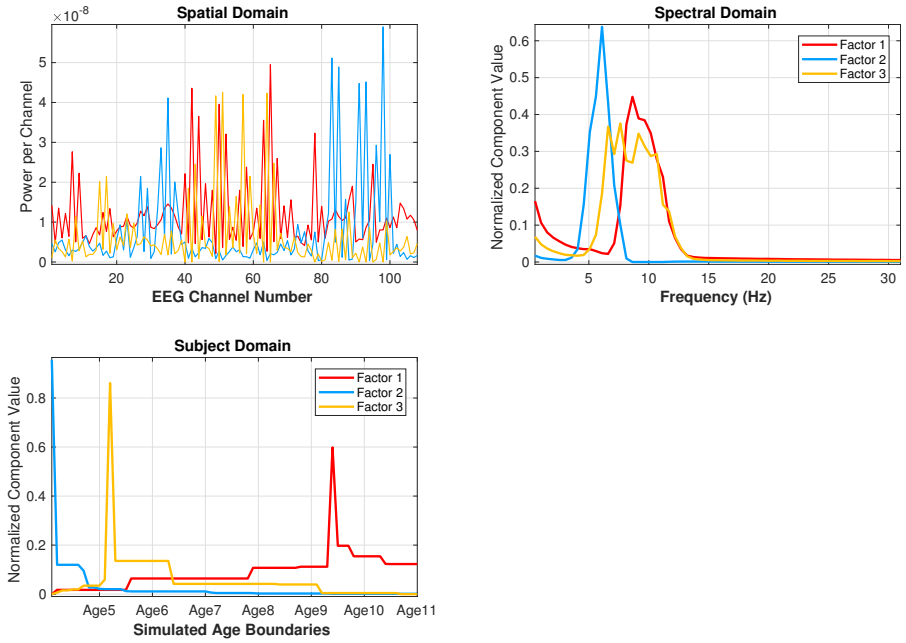

Figure 4.5: The 3-component CP model decomposition of the simulated tensor data. Raw power values are shown for the $[Spatial]$ domain, while component values in the $[Spectral]$ and $[Subject]$ dimension have been normalized. The 3-component model accurately identifies the underlying shifting spectral frequencies set to vary from approximately 5-8 Hz to 8-12 Hz based on age in the $[Subject]$ domain. Sharp peaks in the $[Subject]$ domain reflect the simulated child with the best signal-to-noise ratio.

4.4.1 Simulation

A 3-component CP model revealed the underlying 'ground truth' developmental profiles built into the simulation tensor data. Resulting component factors are presented in Figure 4.5. CORCONDIA was 99% for the model, with approximately 23% explained variance. These values are understandable for the given model, as the BBCB [147] pseudo-EEG simulation was designed to retain trilinearity in data while introducing multiple levels of noise at the frequency band, brain background and sensor layers. Replicating developmental profile extraction results in the simulated data grants further support to our conclusions in real-world datasets.

4.4.2 Preschool Children with Epilepsy: The MMEC Data

RBF-SVM classification of child age was significantly improved using the identified developmental factor profiles as features in the MMEC dataset. Table 4.2 contains the average

Classification	Penalty Cost	Accuracy (%)	Uses CP Features
RBF-SVM	16.0 $\pm$4.1	30.0 $\pm$3.5	Yes
Random Selection	–	20.1 $\pm$0.7**	No
Single Class	25.5 $\pm$1.7†	22.0 $\pm$3.6**	No

Table 4.2: Average classification results across all MMEC cross-validation folds for the RBF-SVM, Random Selection and Single Class classification schemes for epileptic children 0-5 years of age. Features derived from the CP model were only used in training the RBF-SVM classification scheme. The Random Selection and Single Class classification schemes used no information from the CP tensor model. No penalty costs are available for Random Selection classification due to its random nature. No significant differences were found between Random Selection and Single Class classification.
† *Indicates significant difference from the CP trained RBF-SVM using Student's t-test at* $p < 0.05$.
** *Indicates significant difference from the CP trained RBF-SVM based on a one-way ANOVA, corrected for multiple comparisons by Tukey's Honest Significant Difference (HSD) at* $p < 0.01$.

classification results for the MMEC dataset across all training and testing folds for: 1.) RBF-SVM classification based on the estimated direct projection weights for test subjects derived from the CP decomposition features (RBF-SVM trained on CP results); 2.) The RS classification scheme, where random classes were chosen for each test subject (no information used from the CP decomposition); 3.) The SC classification scheme, where all test subjects were assigned to the same class (no information used from the tensor decomposition). The CP trained RBF-SVM classification significantly improved classification accuracy as compared to RS, by approximately 50% (One-way ANOVA, corrected for multiple comparisons by Tukey's HSD $p < 0.01$). Additionally, the RBF-SVM classification was significantly better than the SC classification scheme, by approximately 37% (One-way ANOVA, corrected for multiple comparisons by Tukey's HSD $p < 0.01$). Further, the CP trained RBF-SVM significantly reduced total penalty costs by 9.5 points (37%) compared to the SC classification (Student's *t*-test, $p < 0.05$). To put the reduced penalty cost in context, let's consider a scenario where every test subject has been misclassified. If the SC misclassification was off by approximately 2 years, then the CP RBF-SVM misclassification would only be off by approximately 1.3 years. Average CORCONDIA across training folds was $85.74 \pm 4.86\%$.

While only a third of subjects were on average correctly classified in the MMEC dataset, the significantly reduced cost penalties indicate a move toward less gross age misclassification (e.g. classifying a subject age 0 as age 3, 4, or 5). The improved misclassification penalties indicate

mistaken classification occurring more often for only closely related ages (i.e. children ± 1 or 2 years apart). These improvements along with the factor profiles imply success in identifying developmentally important features of preschool children's EEG.

4.4.3 Child-to-Adult Epilepsy Spectrum: The CHB-MIT Data

Results from the CHB-MIT dataset demonstrate the scalable nature of the proposed analysis across a broader age range than the MMEC dataset. Additionally, subjects are considered to have similar developmental conditions, e.g. epileptic children with matched EEG montages. Due to the limited number of subjects at each age (see Figure 4.1), meaningful classification was not possible for the CHB-MIT dataset. Instead, a qualitative illustration of the general developmental profile trends is presented in Figure 4.6. Importantly, the identified factors reflect expected developmental patterns with respect to which ages are assumed to have the most dominating influence of the $[Spectral]$ domain factors [21, 28]. These results indicate the tensor factorization was able to successfully characterize the developmental features in this dataset. This reinforces the likely generalizable nature of the proposed tensor analysis for use with varied EEG recordings.

Profiles

Figure 4.6 demonstrates the CP model decomposition of the CHB-MIT dataset in a condensed format. Key developmental feature profiles are emphasized across the extracted feature domains, with probable 'background' profiles unaccented. The key feature factors in Figure 4.6 were selected as representative examples which emphasize the underlying development related profiles present in the CHB-MIT data. The key factors were chosen based on their profile dissimilarities to the expected noise profiles in the $[Spectral]$ domain. The feature profiles have been organized by the $[Subject]$ domain, with influential features prominent in early childhood to early adulthood ordered from top to bottom. The exact age of each subject is present on the $[Subject]$ domain axis. Sharp peaks in the $[Spectral]$ domain at 16, 19, 28 Hz are potentially residual artifacts from the time-frequency analysis and NAN-averaging (e.g. ignoring all NAN values) across time-bins. Alternatively, the spikes seen in the $[Spectral]$ domain may be a coping mechanism by the CP decomposition algorithm to account for small variations in the data. The resulting spikes are a result of the CP model optimizing its output,

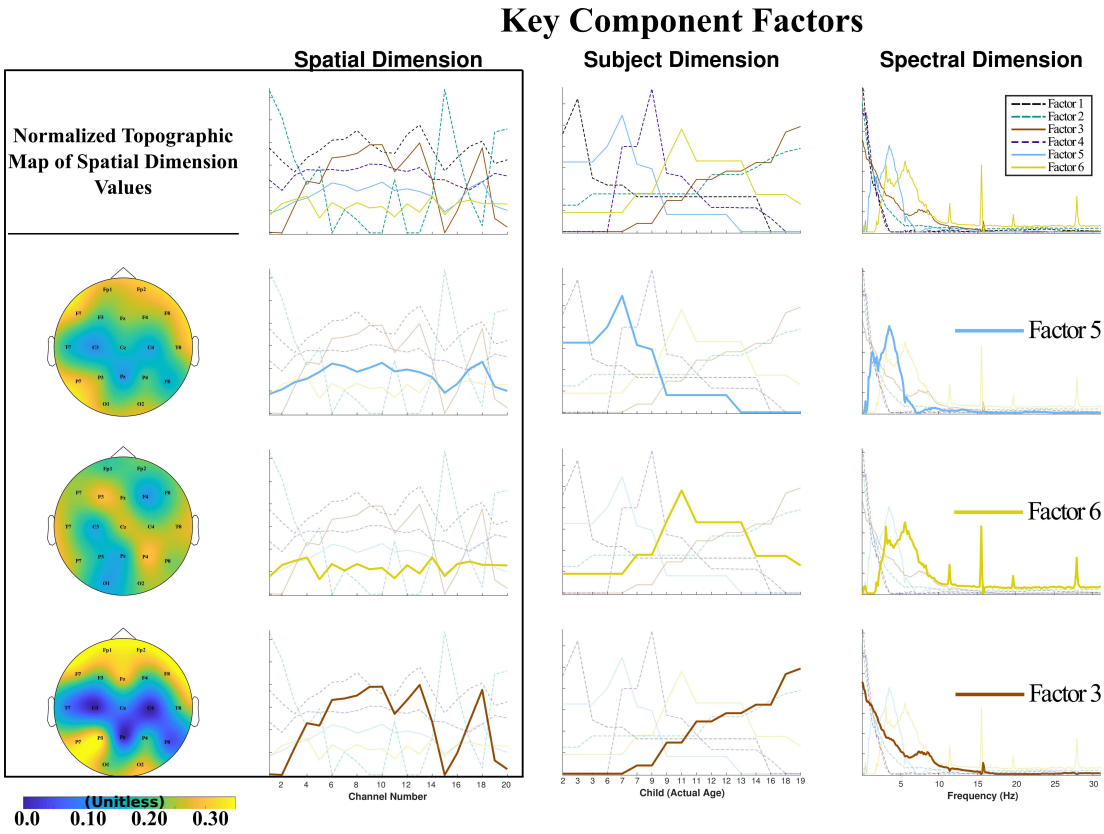

Figure 4.6: A compact visualization of feature profiles from the CHB-MIT dataset for epileptic subjects from age 2 to 19. The key developmental features are separated out for individual inspection, based on their $[Spectral]$ domain profiles. A normalized topographic map reflecting relative contributions of each EEG channel region to the extracted factor is shown in the first column. Combined and separated feature profiles for the $[Spatial] \times [Subject] \times [Spectral]$ domains are shown in columns 2-4.

so reconstruction of the original data using the output component factor matrices is better aligned to the expected values. CORCONDIA was 73.49% for the extracted factors, with 85% explained variance.

4.4.4 School-age Controls: The CMI Data

Developmental features were successfully identified for the healthy control CMI dataset via CP tensor decomposition. Unfortunately, the available control subjects in the CMI data could not be age-matched to the MMEC for direct comparison. However, the features profiles of the CHB-MIT dataset (seen in Figure 4.6) demonstrated maturational differences could be found between older children, including those within the age range of the CMI subjects.

Thus feasibility of finding features sensitive to development in (relatively) older children was supported. Table 4.3 contains the average classification results for CMI subjects. Again, the three classification schemes of RBF-SVM, RS and SC were investigated. Results from the CP decomposition were only used in training the RBF-SVM. As such, the CP trained RBF-SVM classification significantly improved classification accuracy compared to RS, by approximately 25% (one-way ANOVA, corrected for multiple comparisons by Tukey's HSD $p < 0.01$). Classification accuracy between the RBF-SVM and SC schemes was also found to be significantly different (one-way ANOVA, corrected for multiple comparisons by Tukey's HSD $p < 0.05$). The RBF-SVM classification significantly reduced total penalty costs by 2 points (19%) compared to SC classification (Student's *t*-test, $p < 0.05$). Average CORCONDIA across training folds was $78.25 \pm 6.57\%$.

Although overall improvements in classification results for the CMI were smaller than those in the MMEC dataset, significant improvements were still uncovered. Extra difficulty in discerning developmental differences in the CMI data is not surprising. There is a smaller developmental window spanning subjects in the CMI dataset. Additionally, CMI subjects were likely more homogeneous compared to the epileptic population in the MMEC. Therefore, despite only marginal boons in accuracy and penalty costs, the classification improvements in the CMI dataset under CP modelling are important. The positive results serve as evidence that the proposed tensor analysis is accessible to typical developing children.

The CMI dataset results also further illustrates the scalable nature of the tensor analysis. The findings demonstrate a successful feature extraction on subjects more similar developmentally (i.e. healthy, between the ages of 8-11 y.o.) compared to the other datasets. Physiological changes across this age span are significantly less drastic compared to both the MMEC development window (from infancy to early childhood; age 0-5 y.o.) and the CHB-MIT development window (from infancy to early adult hood; age 2-19 y.o.).

4.4.5 t-SNE Visualization

Features from the first training fold tensor decomposition of the MMEC and CMI datasets are displayed as t-SNE maps in Figure 4.7(b) and Figure 4.7(d) respectively. Both t-SNE maps demonstrate strong local grouping of different age groups. Maps of EEG data prior

Classification	Penalty Cost	Accuracy (%)	Uses CP Features
RBF-SVM	9.6 $\pm$0.9	34.2 $\pm$1.9	Yes
Random Selection	–	27.4 $\pm$1.3**	No
Single Class	11.8 $\pm$1.1†	29.8 $\pm$3.6*	No

Table 4.3: Average classification results across all CMI cross-validation folds for the RBF-SVM, Random Selection and Single Class classification schemes for healthy children 8-11 years of age. Features from the CP model were only used in training the RBF-SVM classifier. The Random Selection and Single Class classifiers used no information from the CP tensor model. No penalty costs are available for random classification due to its random nature. No significant differences were found between Random Selection and Single Class classification.
† *Indicates significant difference from the CP trained RBF-SVM using Student's t-test at* $p < 0.05$.
* *Indicates significant difference from the CP trained RBF-SVM based on a one-way ANOVA, corrected for multiple comparisons by Tukey's Honest Significant Difference (HSD), at* $p < 0.05$.
** *Indicates significant difference from the CP trained RBF-SVM using a one-way ANOVA, corrected for multiple comparisons by Tukey's Honest Significant Difference (HSD), at* $p < 0.01$.

to tensor factorization (Figure 4.7(a)) and when the decomposition has a randomly ordered $[Subject]$ domain in the MMEC data (Figure 4.7(c)) are included for comparison. The results demonstrate significantly improved feature grouping and clusters using the properly ordered tensor decomposition methodology compared to using the raw EEG or random ordered $[Subject]$ domain. The t-SNE maps illustrate that training the RBF-SVM with raw EEG time-series data, e.g. using all frequency and spatial features, does not perform better than just random classification in the MMEC dataset. Similar patterned results are found for the CMI data.

4.5 Discussion

The tensor analysis outlined in this chapter lays a foundational framework capable of extracting latent structures and features associated with development in paediatric EEG. The unsupervised nature of this framework opens the door to encourage better personalized paradigms for data-driven technologies aimed at paediatric populations. Capitalizing on these developments could help translate new technologies sensitive to developmental features in EEG, like BCI, to children.

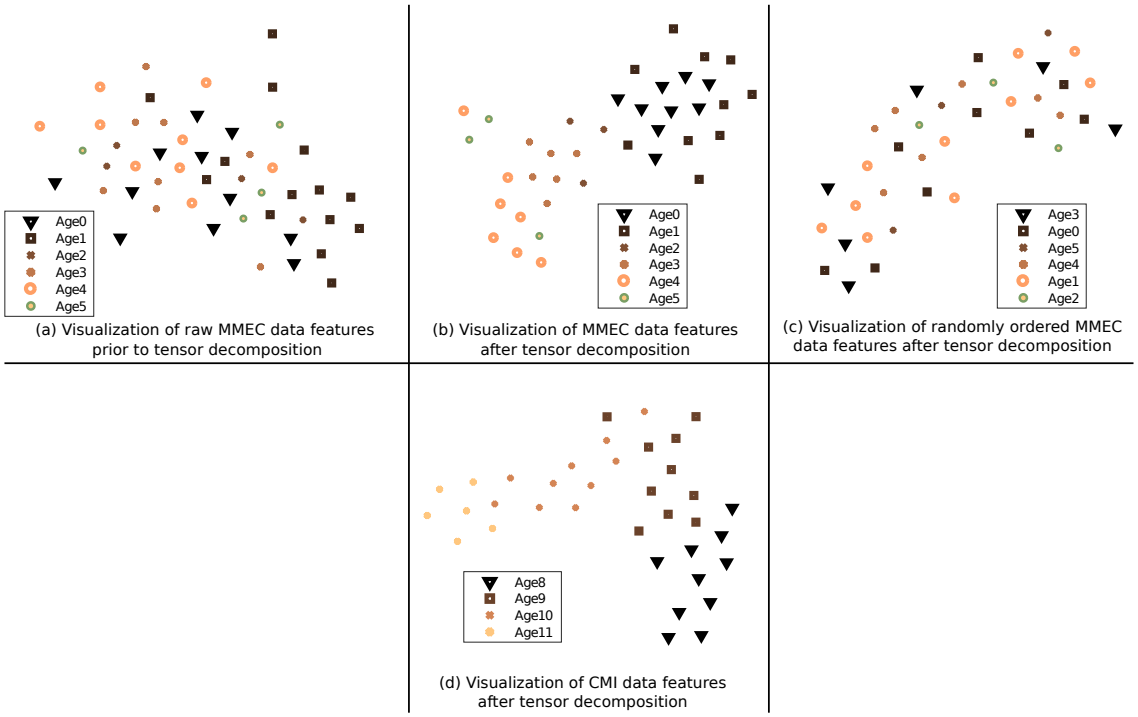

Figure 4.7: t-SNE Map visualizations with features grouped by subject age. (a) Map of EEG data prior to tensor decomposition of the MMEC dataset. (b) Map of CP extracted developmental features of the MMEC dataset, using the first cross-validaiton training fold. (c) Map of CP extracted developmental features of the MMEC dataset when the original tensor has a randomly ordered $[Subject]$ domain. (d) Map of CP extracted features from the CMI dataset using the first cross-validation training fold.

4.5.1 Tensor Analysis Derives Informative 'Development Feature Profiles' of Paediatric Subjects

The factor profiles derived in this study confirm that the proposed CP decomposition simultaneously accounts for various EEG brain processes and shifting frequency bands across subject age. The presented results often explained more than 85% of the data variance. Low-frequency, high power spectral activity typically associated with background EEG physiological processes can be seen in Figure 4.4. Factors 1-4 reflect typical noise decaying power curves in the $[Spectral]$ domain, coupled with relatively strong contributions across all channels in the $[Spatial]$ domain. These factors are likely characterizing general irrelevant brain processes of the EEG, e.g. the electrophysiological processes generated by involuntary systems of the body, independent to potential signals of interest (e.g. factors 5-8). Commonly, this may be referred to as 'background physiological EEG noise'. Critically, the $[Subject]$ domain demonstrates the tensor analysis has sensitivity to subtle developmental differences, even within the potential background processes. This is supported by the fact that the likely background processing features are captured in a manner which is also uniquely corresponding to distinct age groups.

Shifts in spectral power and frequency due to development [12,21,22] are seen in the qualitative factor profiles presented for both the MMEC and CHB-MIT datasets. In the MMEC dataset, for example, the very low frequency, high power dominated spectral profile associated with infant and early-life EEG recordings seem to be reflected by factor 7 in the $[Subject]$ and $[Spectral]$ domains. Meanwhile, factor 5 (light blue) is centrally located in the $[Subject]$ domain (covering approximately ages 1-3) and spans the 3-7 Hz range of the $[Spectral]$ domain. Factor 5 therefore highlights the likely dominant frequency range for those ages. Factor 5 also reflects a shift in power towards higher frequencies, which is expected with maturation. Factor 8 illustrates a further shift towards higher frequencies contributing more to the spectral profile, as it steadily increases for subjects 31-50 (approximately ages 3-5 y.o.) in the $[Subject]$ domain. Also, factors 6 and 8 may represent the beginning separation between classical EEG bands of interest, i.e. the delta/theta bands and the alpha/beta bands respectively.

Similar developmental shifts are also seen in the CHB-MIT qualitative feature profiles. The

key component factors show significantly reduced spectral power contributing more in higher $[Subject]$ domain ages. The key factors also demonstrate a shift in the $[Spectral]$ dimension towards higher frequencies. These extracted profiles appear to reflect the traditional movement and prevalence of the classical EEG bands, like alpha, throughout child development into adulthood [12, 21].

4.5.2 Improved classification results verify tensor extracted features' sensitivity to development

Age-specific factors uncovered from the EEG via CP decomposition contain structural information on latent developmental relationships. Improvements in the CP trained classification coupled with the feature profiles obtained from the tensor analysis for both impaired (epileptic) and healthy children supports this finding. The scalable nature of the proposed tensor framework is promising. The analysis was shown to be capable of identifying relevant developmental features across varying conditions including both afflicted/healthy populations, and at slow/rapid developmental windows.

Better characterization of such features for paediatric subjects invites a stronger case for translating signal processing and machine learning applications to children. Clear support for this is seen in the comparison between t-SNE maps for the full feature raw EEG time-series data and the CP processed data. Using the feature-full raw EEG time-series without processing led to complete failure in identifying developmental features. The resulting t-SNE map has no discernible structure with random clusters and groupings. Characterization of the underlying developmental profile was rendered completely imperceptible, which is likely reflected in the classification being no better than chance. On the other hand, the highly-structured t-SNE maps of the tensor extracted features for both the MMEC and CMI dataset reflect well characterized developmental profiles. Such profiles could be utilized in machine learning applications.

Importantly, the randomly ordered $[Subject]$ domain t-SNE map suggests that successful identification of key developmental features is not inherent to tensor factorization itself. Although the overarching global structure is capable of being identified, as seen in the similarity of global shapes of Figure 4.7(b) and Figure4.7(c), the local grouping is completely lost. Rather, the viability of determining developmentally sensitive features relies on proper

construction of a developmental measure, e.g. the $[Subject]$ domain in this study. The use of a strictly increasing subject age, together with the unimodality constraint, acts as a natural proxy to measure maturation in this work.

Together these findings support our proposed tensor analysis framework. As discussed in Section 2.5.6, interpretation of the results relies on understanding the relationship between the data evaluated. Our methodology is verified using multiple datasets. From this, the multi-way analysis model can be seen as capable of handling various types of cross-sectional developmental data. The results indicate the new proposed methods can account for developmental differences in background EEG processes and shifting spectral signals for children under a variety of different developmental conditions. Thus, the proposed analysis can capture latent information concerning the relationship to how a child is developing, even in situations where the involuntary brain processes are masking the child's signals of interest. By being able to decipher differences in the background EEG processes, the proposed analysis further provides a potential tool in evaluating potential differences for the child's development with respect to brain processes which may be less directly measurable through cognitive tasks (e.g. attention). The classification results illustrate a means for developmental feature extraction sensitive to progressive changes. As such, classification in this manner provided proof-of-concept for being able to extract specific cognitive features of interest from a distributed set of developmental states. Furthermore, the identified feature profiles provide informative context regarding the relationship between $[Spatial]$ and $[Spectral]$ structures relative to subject age and development. These findings provide a means for research scientists and clinicians alike to probe the changes in relationship between these modes for distinct developing subsets of the paediatric population. By being able to characterize typical development and the expected relationship between the $[Spatial]$ and $[Spectral]$ structures of the EEG, deviations from this expected relationship may be useful in understanding differences in a child's development. Furthermore, tensor factorization is shown to organize into locally relevant developmental features when proper consideration is given to the representation of maturation in a structured domain. Thus cross-sectional analysis could be developed using multi-way methodology to establish a set of relationships present in typical developing paediatric EEG data, so children who fall outside these expected relationships are readily identified from a given population.

The findings of this chapter outline a methodological framework engineered to natively identify and extract EEG signal features related to development.

4.5.3 Potential Applications and Future Work

The ability to exploit inherent structural information from higher-order tensors offers potential new foundations for designing EEG-based applications sensitive to maturational differences. The proposed methodology could therefore be useful in translating technology like BCI to children. Accordingly, adapting the proposed tensor analysis outlined in this chapter to a BCI-like paradigm was the motivation for Chapter 5. Incorporating other developmental measures, like cognitive or behavioural scores, into the validated tensor structure is investigated in Chapter 6, to broaden the scope of developmental behavior captured by the tensor framework.

Future work could include expanding the proposed methodological framework to determine a current child's developmental state at an electrophysiological level. This could help tune the BCI appropriately for subjects at different developmental points. Additionally, cataloguing 'healthy development curves' could be explored to permit tracking development in impaired children is another potential option for expanding this research. Furthermore the misclassification costs presented only consider a linear cost matrix for discerning penalties between misclassified ages. Future work could examine other alternatives to the linear cost matrix for better representation of the relationship between misclassified ages.

Additionally, using the numerical values relating to children's age in the $[Subject]$ domain of the multi-way analysis may have an unknown effect on the model results. Therefore developing a neutral codification of child age, rather than the raw numeric value, for the $[Subject]$ domain could be of interest. Using such a codification scheme, potential biases resulting from using numeric values would be mitigated. This is a promising avenue for future work.

4.5.4 Limitations

Limitations in this study included restricted access to age/acquisition-matched paediatric datasets and the heterogeneity associated with epilepsy in subjects. Due to limited resources, no direct age/acquisition-matched healthy control data was available for analysis to compare

directly. To mitigate these drawbacks, however, multiple publicly available datasets were used to demonstrate the proposed methodology in multiple settings. The CHB-MIT dataset built upon our analysis from the MMEC across a wider age range using a similar disease condition, while the CMI dataset represented healthy control within the bounds of childhood (but not age-matched). Future work using a more homogeneous population with age-matched controls could help further validate the results, along with data from both resting-state and event-related EEG.

4.5.5 Pilot Study: Replicating Analysis in Paediatric Stroke Data

Preliminary results from a small collaboration grant with the Alberta Children's Hospital provided an opportunity to replicate the framework developed above for a paediatric stroke population. The pilot study aimed to replicate the rEEG analysis using tensors for a paediatirc population who could readily benefit from BCI implementation, e.g. children who had suffered from arterial ischemic stroke (AIS) early in life. These results are presented in Appendix B and lend further support to the conclusions presented in this chapter.

4.6 Conclusion

Advanced signal processing, like CP tensor decomposition, combined with machine learning can help distinguish non-obvious developmental patterns from child EEG data. This chapter demonstrates tensor analysis can provide an intuitive sense of the latent developmental relationships in paediatric EEG data and supplies a natural means for development-sensitive feature extraction. The reported results indicate successful identification of feature profiles from a variety of development conditions, including both healthy and afflicted paediatric populations. The tensor analysis framework is shown to improve classification and generate local and global structures in paediatric EEG data. Thus, this study lays a methodological framework which could improve paediatric applications reliant on EEG signal processing and analysis.

Chapter 5

Extensions to 4-Mode Tensors and SSVEP Data

This chapter is based upon the accepted conference paper for the EUSIPCO 2018 conference [4].

The key contributions of this chapter are:

- Demonstrating how to extend the multi-way framework to an event-triggered EEG signal, e.g. steady-state visual evoked potentials.
- Utilize orthogonality and non-negativity constraints in the CP model to develop an SSVEP binary classification paradigm which does not use subject-specific calibration.
- Show the tensor framework's capability to predict subject data within a BCI setting.

5.1 Introduction

A common brain signal exploited in BCI technology is the steady-state visual evoked potential (SSVEP) [30, 31, 41, 135]. The SSVEP is a natural response in the brain which occurs when visual stimulation is presented at specific lower range frequencies, e.g. below 100 Hz [31]. The 'steady-state' aspect refers to a continued exposure to an oscillating visual stimulus, which then elicits a stable, evoked potential identifiable in EEG recordings over the occipital (visual) region of the brain [30, 31, 135]. When a person attends to the stimulus, the natural brain rhythms over the occipital region synchronize to match the stimulus' oscillating frequency, and the oscillations can be detected via the scalp EEG electrodes [31, 135]. This phenomena is associated with spatial attention, and as such the source of the SSVEP-stimulus can be matched

to specific choices, e.g. letters on a computer screen or a virtual button [30, 31, 135]. BCI paradigms utilize the SSVEP by having the user attend to a desired target in order to 'click' the button or 'write' the letter [30, 31, 135].

The predictable nature of SSVEP signals make them enticing for building BCI paradigms for complex populations, such as impaired children. However, the scope of structural and functional changes associated with child development and the challenges discussed in Chapter 3 for BCI applications in children are still present for SSVEP-based BCIs. Previous work on SSVEP-based paediatric BCIs have demonstrated inherent age-specific responses affecting SSVEP-based BCI tasks [31, 115, 135]. Results have also indicated a link exists between the brain's ability to synchronize spontaneous and steady-state evoked oscillations with its maturation throughout childhood [115, 148]. In principal, this is once again related to the common hurdle facing paediatric BCIs, e.g. the ongoing progression in electrophysiological signals and networks recorded by EEG changing throughout development [22, 28]. Therefore, a potential avenue to incorporate these maturational differences within EEG signals is by extending our proposed multi-way tensor analysis introduced in Chapter 4 to the SSVEP-based stimulus data, with the aim to maintain and investigate the higher-order structures of SSVEP-driven EEG across developing subjects.

Exploiting advantages gained by tensor analysis has been previously demonstrated as an effective tool in SSVEP-based BCIs for adults [41, 122, 149]. Rapid, accurate detection and classification of brain signals is crucial for real-time analysis in SSVEP-based applications and thus can benefit from tensor-derived characteristic filters, as outlined in Chapter 2. Additionally, tensor-based BCI results have established tensor decompositions as a potential tool for avoiding the subject-specific calibration in BCI classification [150]. Reducing the calibration time for a subject can lead to improved BCI set-up speed and increases its child-friendly nature, as described in Chapter 3. However, little progress has been made in adapting SSVEP-based BCI paradigms to children [31, 77, 135], especially with respect to improved signal processing of EEG. Therefore, this chapter outlines for the first time initial steps in adjusting SSVEP signal processing techniques common in BCI paradigms to accommodate paediatric populations. The work proposes an effective tensor analysis scheme to classify binary SSVEP selection tasks in children, and demonstrates its potential using a

publicly available dataset.

5.2 Materials

5.2.1 Child Mind Institute Dataset

Data in this study was derived from publicly available EEG provided by the Child Mind Institute (CMI) [25]. EEG data was recorded from a high-density 129-electrode hydro-cel EEG cap at a sampling rate of 500 Hz, and filtered by the CMI at a bandpass of [0.1 - 100] Hz. Data selected for this chapter again included only pre-adolescent subjects from the CMI dataset ranging from 6 to 11 years of age [25]. Each child was subjected to a suite of task-independent (passive) and task-based (active) paradigms designed to capture a comprehensive set of EEG recordings complementary to investigations of human brain function [25]. Participants were seated in a sound-attenuated, dark experiment room approximately 70 cm from a 17-inch CRT monitor, with a vertical refresh rate of 100 Hz [25]. Subjects were instructed to stay as still as possible during the tasks, and given two breaks throughout the EEG session during which impedance levels of electrodes were checked and reduced as necessary [25]. Stimulus and other paradigm presentations were programmed in Matlab using various extensions (see [25] for details), with each child receiving the EEG tasks in the same order. Compliance for completing the experiments was proctored by a research assistant, present in the room for each of the subjects considered in this analysis [25]. Multiple recordings were taken for each child, spanning across the task-independent and task-oriented paradigms. Both raw and pre-processed EEG data was available for analysis.

Contrast Change Paradigm Overview

From the various EEG tasks provided by the CMI, the *Contrast Change Paradigm* (CCP) was selected for analysis, as it included a means to detect isolated SSVEP-driven EEG data. The CCP task design was based upon a method used to isolate and simultaneously trace the neural dynamics of several key processing stages underlying simple sensorimotor decisions [25, 151]. These key processing stages include sensory evidence encoding, evidence accumulation over time, and motor preparation [151]. The CCP design modified this task to include the ability to

probe fluctuations in attentional engagement in addition to the other sensory-motor processing levels [25]. An illustration of the CCP is given in Figure 5.1.

For this chapter, particular interest was given to the sensory evidence encoding stage as it includes SSVEP-driven EEG signals. The CCP used a continuously presented, oscillating visual stimulus with gradual-change in targets for the subjects to eliminate the transient, task-irrelevant and sensory-evoked signals which often accompany discrete EEG task stimulus [25]. As a consequence, the continuous visual stimulus in the CCP produces measurable SSVEP in the subject during this phase. In practice, the CCP employed two simultaneous stimuli with gradual changes on a left and right object, requiring the user to monitor changes in terms of contrast between the two flickering objects [25] and thereby functioned as an off-line BCI task. The two oscillating objects started at equal contrast (50%) in the baseline (e.g. between targets), with one gradually increasing to 100% and the other simultaneously decreasing to 0% over the course of the trial [25]. The changes in contrast occurred linearly over 1600 ms, with an immediate 800 ms linear return to 50% following the end of the trial [25]. The left and right objects oscillated at 20 and 25 Hz, respectively [25]. Subjects were instructed to attend to which object had its contrast increasing, and push the corresponding left or right button [25]. A random set of 24 distinct trials (12 left, 12 right) were run in a 1-minute block of the CCP, and each child attempted to complete 3 CCP blocks. Of the initial 45 subjects available, $n = 40$ children were able to complete at least 2 of the 3 blocks with good responses, and were therefore selected for classification analysis in this chapter.

5.3 Methods

5.3.1 Pre-processing

With respect to cleaning the raw EEG for analysis, we opted to use the pre-processed EEG data provided by the CMI. This helped to improve repeatability in the proposed analysis, by limiting potential variances in the EEG pre-processing pipeline, such as artifact detection and removal. This was handled at the point of distribution by the CMI, and is guaranteed to be identical for any replication experiments.

Isolating the SSVEP signals from the CCP required some additional processing. The given data

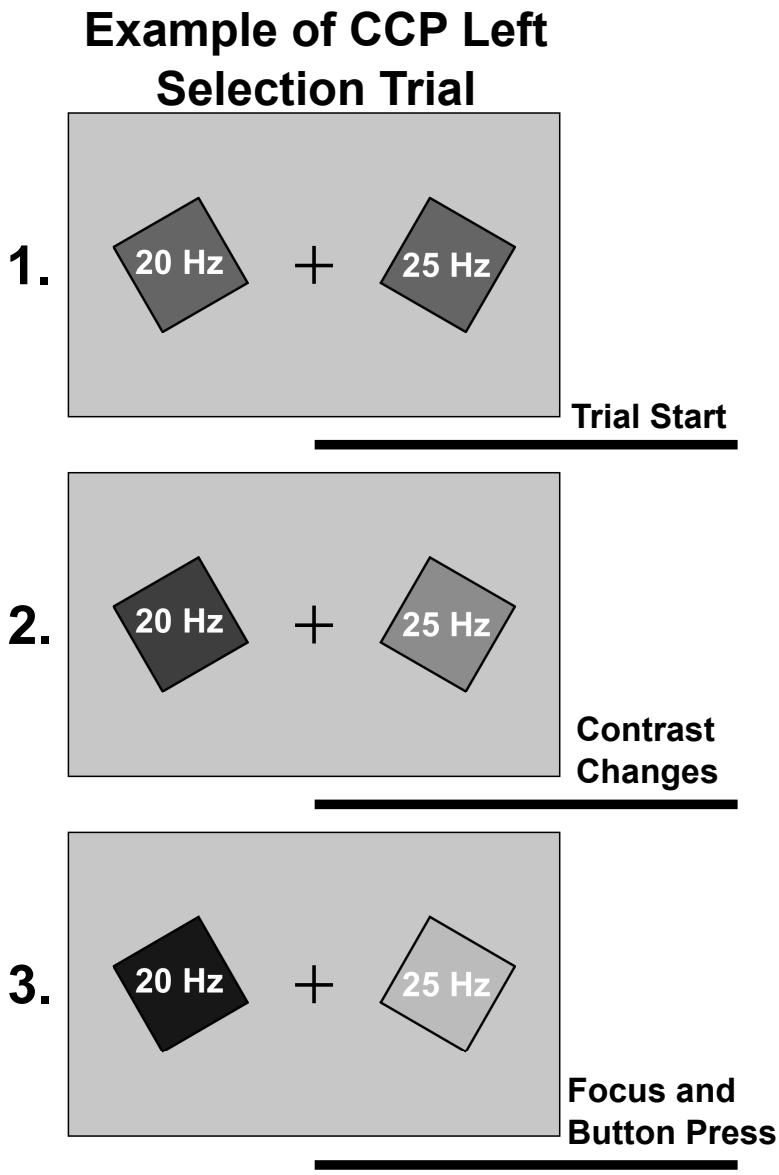

Figure 5.1: A visual representation of a CCP trial, for a left object selection. Panel 1. presents the baseline screen of the CCP, with both the left and right objects presented at 50% contrast to the subject on a computer screen. Panel 2. illustrates after the trial has begun, showing the left object increasing in contrast, and the right object decreasing in contrast. Panel 3. shows when the contrast change is clear enough for the patient to focus on the left (20 Hz) object near 100% contrast, and press the corresponding button.

was re-referenced to a common average reference (CAR), matching the analysis conventions established for the rEEG analysis in Chapter 4. For each subject, the continuous EEG recording was segmented into non-overlapping trials based on provided event flags in the EEG data. A trial was defined as the 2 seconds prior to a button press, enabling the analysis to fully encompass the contrast change. Spectral power in the frequency range 17-30Hz was calculated for each trial across 18 occipital channels using the Fieldtrip toolbox for Matlab [58]. Trials which had poor EEG channel interpolation and/or abnormally high amplitude variance (i.e. a 10-fold difference) were excluded from analysis. The 'ground truth' SSVEP data was labelled as either left or right, based on the corresponding button press.

5.3.2 Tensor Construction, Decomposition and Modelling

From the clean SSVEP data, a 4-mode tensor $\mathcal{X}$ was built with domains $[Trial] \times [Spatial] \times [Spectral] \times [Subject]$. The $[Trial]$ domain was structured such that the first 12 elements were left-trials, followed by 12 right-trials. This removed the randomized left/right trial order in any given CCP block, allowing for the $[Subject]$ domain to be compared in the constructed tensor. Furthermore, the $[Subject]$ domain was arranged from youngest to oldest to account for potential developmental differences associated with age. The tensor construction was therefore similar in practice to the rEEG tensors developed in Chapter 4, but with the added $[Trial]$ domain. Figure 5.2 illustrates the described 4-mode tensor.

The constructed tensor $\mathcal{X}$ was then decomposed into a 3-component CP model [42,46], similar to Chapter 4. The CP model was analyzed using the N-way and Tensorlab toolboxes for Matlab [59, 60]. The CP decomposition model was again chosen to represent $\mathcal{X}$ as the sum of component rank-one tensors, thus exploiting the 1:1 interaction between components across the 4 modes of the tensor. Equation 5.1 shows a three component decomposition for our 4-mode tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K \times N}$ with component matrices $\mathbf{A} \in \mathbb{R}^{I \times R}$, $\mathbf{B} \in \mathbb{R}^{J \times R}$, $\mathbf{C} \in \mathbb{R}^{K \times R}$, $\mathbf{D} \in \mathbb{R}^{N \times R}$ where $R = 3$.

$$\mathcal{X} = \sum_{r=1}^{R=3} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r \circ \mathbf{d}_r \tag{5.1}$$

Similar to Chapter 4 and analysis of the rEEG data, non-negativity was enforced in the $[Trial]$,

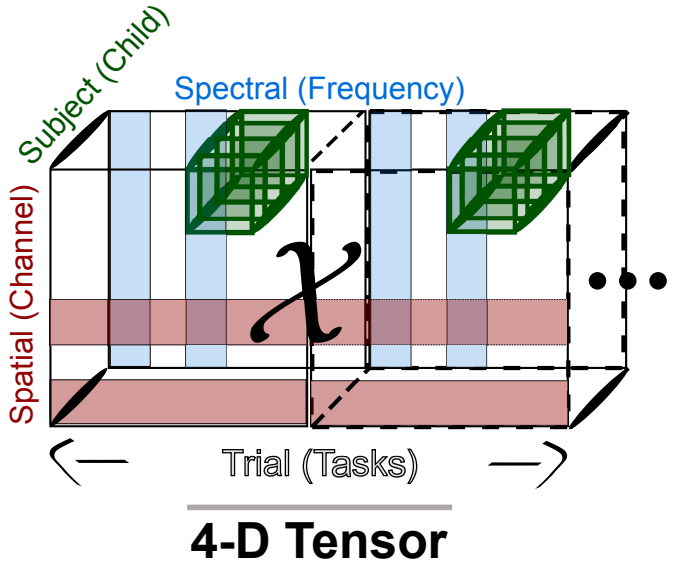

Figure 5.2: Graphical representation of the extended 3-way tensor to include a 4th mode, the $[Trial]$ domain, to account for the left and right selection from the SSVEP task. Each separate cube represents the set of $[Spatial]$, $[Spectral]$ and $[Subject]$ mode information corresponding to a given trial. The ellipsis indicates a continuation of the structure for as many individual trials and/or tasks present in the data.

$[Spatial]$ and $[Subject]$ domains. The non-negativity constraint improved interpretation of the resulting component matrices, as discussed in Section 4.3.3. In addition, orthogonality was imposed on the $[Spectral]$ domain. Orthogonality was introduced to guarantee linear independence between factors, with the goal of representing the distinct frequency SSVEP peaks at 20 and 25 Hz as independent sources. While orthogonality reduced the interpretation of the spectral information to some degree, (i.e. negative spectral values are possible under an orthogonal constraint), the ubiquitous and dominating properties of the SSVEP signal should be distinct enough to mitigate this concern. The dominating frequency peaks at 20 and 25 Hz in the bottom left panel of Figure 5.3 of the results help support this conclusion.

A 3-component decomposition was selected so distinct factors for the two SSVEP signals of interest could be uniquely described, with the third factor accounting for any other variability or irrelevant noise. Using fewer than 3 components would mean the CP model could not separate the desired signals distinctly from each other and noise. Using more than 3 components in the decomposition was examined, but found unsuitable as mentioned in the Results section below.

5.3.3 Projection and Classification

For each CCP-block of data, the whole block was designated as either 'training' or 'testing'. It is important to note that training and testing here do not refer to a classifier, as seen in Chapter 4, but rather are used to designate blocks which are decomposed using the CP tensor model (training), and which are the unseen data for validation (testing). Of the $n = 40$ children identified in the pre-processing, a unique subset of $n = 34, 36, 35$ children were found to have successfully completed CCP-blocks 1, 2 and 3, respectively. Tensor decomposition and the direct projection validation was performed across the six combinations of the 3-block CCP data, e.g. Block1-Train, Block2-Test; Block1-Train, Block3-Test; etc.

Training Blocks

In training blocks, the 1-minute CCP-block of data was decomposed into its relevant 3-component matrices via a CP decomposition. These component matrices were then used to estimate new weights for the unseen data in the testing block using direct projection and characteristic filters as outlined in Chapter 2. Importantly, we opted to estimate the weights of the new data for each trial independently, as expected in a real BCI scenario. This was similar in approach to work done in [150]. Information from the $[Trial]$ domain could not be explicitly included in the encoding matrix $\mathbf{W}$, so it resembled the 3-mode example given in Chapter 2 (e.g. equations (2.20 – 2.23)).

This occurs in part because of the direct projection steps (2.22) and (2.23), where the unseen data is flattened and projected onto the encoding matrix of $\mathbf{W}$. The multiplication requires an exact matching number of elements between the projection domain and the unseen data. However, if the $[Trial]$ domain is included in its entirety, but only one new trial is being evaluated (as would be expected in a real-time BCI application), there would be a 1:24 element ratio between the two $[Trial]$ domains. Therefore the encoding matrix $\mathbf{W}$ in this analysis is comprised of only the $[Spatial]$ and $[Spectral]$ weights. Critically, however, with the strict nature of the super-diagonal core in a CP decomposition, the estimated interactions in $\mathbf{W}$ have inherited properties corresponding to $[Trial]$ (and $[Subject]$) domain structures, as outlined in Figure 5.3. Therefore no important structural information is lost when estimating the new data using $\mathbf{W}$.

Testing Blocks

The testing block was comprised of solely unseen data not used in the CP model decomposition. Within the block, trials for every subject were selected randomly for validation using the direct projection method. New subjects from the testing tensor $\mathcal{Y}$ were projected onto the common interactions held by $\mathbf{W}$ to estimate weights w_y for the unseen data, as described above. These weights informed on the new subject's status with respect to the extracted tensor components. Knowing which components in the CP model reflected the noise, left and right SSVEP signals (as seen in the bottom-left $[Spectral]$ domain panel of Figure 5.3) permitted each given trial to be classified as 'left' or 'right' based on comparing the corresponding component weights from w_y. Ground truth was taken to be the actual button pressed for the trial by the subject.

5.4 Results

The 3-component tensor decomposition of each CCP-block provided a set of component matrices suitable for the classifier-free training and testing. Explained variance for the constrained 3-component model was approximately 41% as compared to the 43% explained variance without constraints. Using more components for decomposition did not significantly improve the explained variance of the model (5-components had 42% explained variance), and introduced a greater risk of overfitting in the $[Subject]$ domain. Visual inspection of the components in the $[Trial]$ and $[Spectral]$ domains provided easy identification of which component corresponded to noise, left trial and right trials. Figure 5.3 shows the 3-component CP decomposition of the first 1-minute block in the CCP data.

Left, right and total accuracy were calculated for each child based on the mean across all trial classifications estimated by the direct projection. Figure 5.4 shows a box-plot of resulting accuracy across all children and all trials for each CCP-block combination. Classification above 66.7% is considered to be significantly greater than chance at $p = 0.05$ for a binomial distribution of 24 trials with equal classification probability.

Figure 5.5 illustrates the raw power spectrum for occipital channels in two representative children during left-selection SSVEP (20 Hz) tasks. The results include a random subset of 6 individual trials and an 'overall mean' comparison between prototypical 'good SSVEP' and

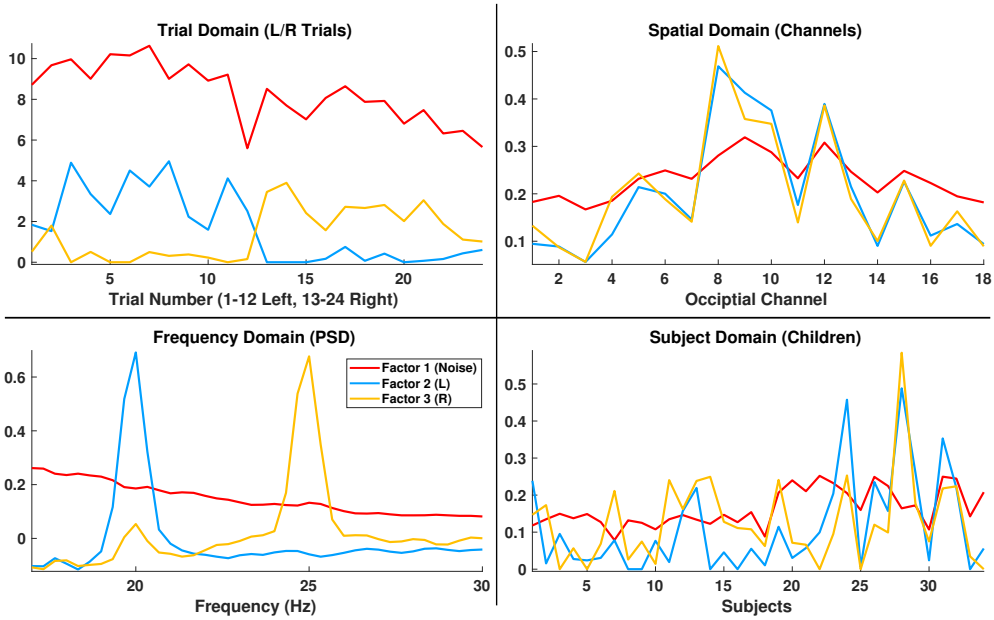

Figure 5.3: Results from the constrained 3-component CP decomposition of the 4-mode SSVEP data in CCP block 1. Across component domains, interaction occurs only between the factors of the same color, e.g. red relating to red across all panels. Here, the $[Spectral]$ domain captures background noise and target SSVEP signals at 20, 25 Hz in the red, blue and yellow factors respectively. Similar decompositions are found for CCP blocks 2 and 3.

'poor SSVEP' response children.

The results indicate the extracted tensor components are relatively robust for classification, even when the target signal is not immediately clear, like in Child 2 of Figure 5.5. For the example children, the proposed tensor classification scheme was 100% accurate in predicting any given trial as 'left' in Child 1, and 64% accurate for Child 2. By utilizing information from both $[Trial]$ and $[Subject]$ domains in the original training tensor, the CP model decomposition can implicitly emphasize which trials (and subjects) are representative (or not) of the signals of interest. This is evident when comparing the loading contributions for the left-selection (20 Hz) signal of Child 1 and Child 2 in the $[Subject]$ domain panel of Figure 5.3. Child 1 contributes approximately 15x more to defining the extracted left-selection component compared to Child 2 ($l = 0.24$ for Child 1, $l = 0.016$ for Child 2). Despite this large disparity, left-selection trials from Child 2 were still correctly identified approximately 64% of the time. This indicates that accuracy rates for signals of interest can largely be maintained for noisy, unclear trials.

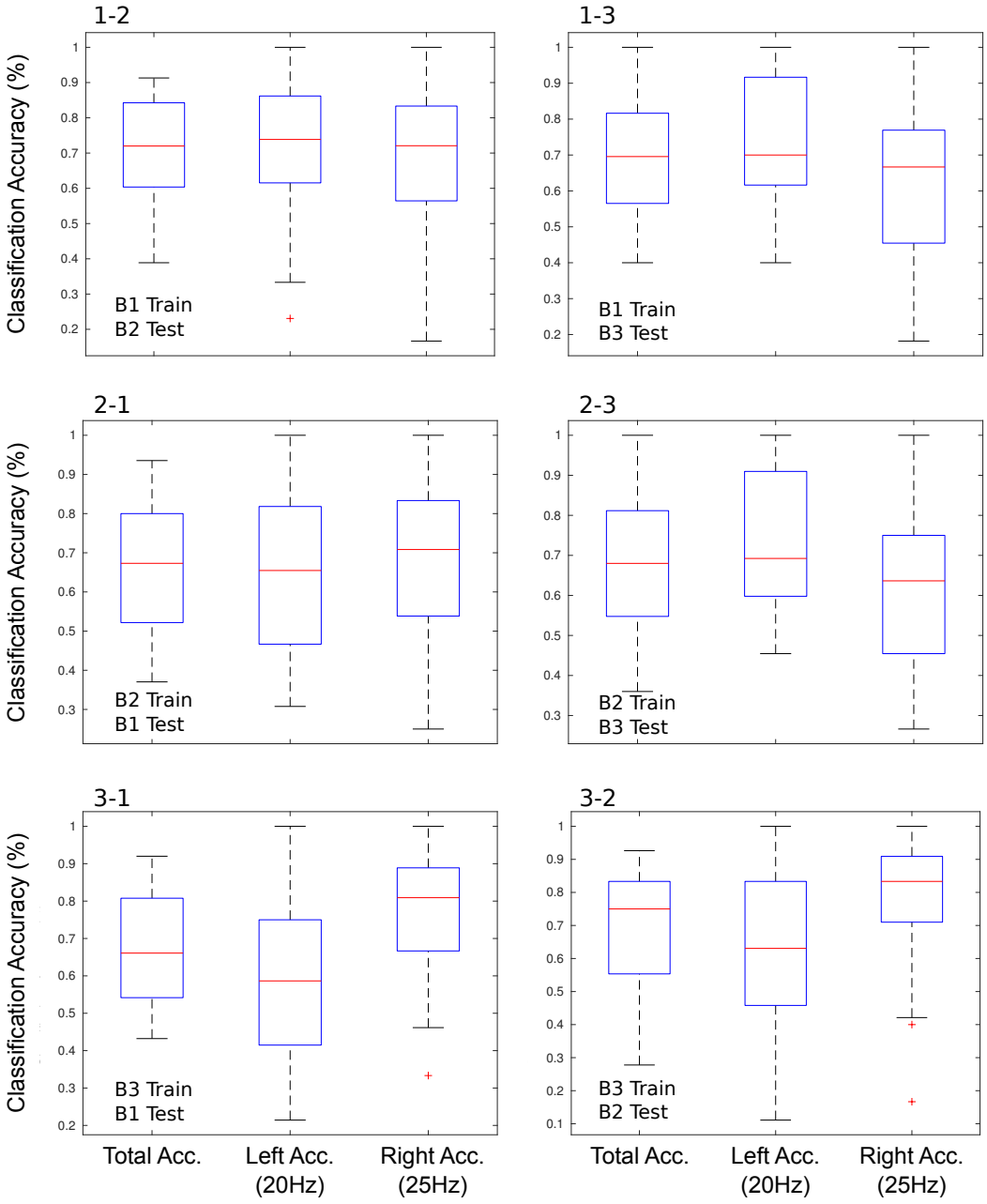

Figure 5.4: Box-plot results of the average total, left-target and right-target accuracy across subjects and trials. Each panel shows individual train-test block accuracy across all children who completed the given CCP block ($n = 34, 36, 35$ for blocks 1, 2 and 3, respectively). Total accuracy was relatively consistent across CCP combinations, at approximately 70%. Outlier subjects are indicated beyond the box-plot range by a + symbol.

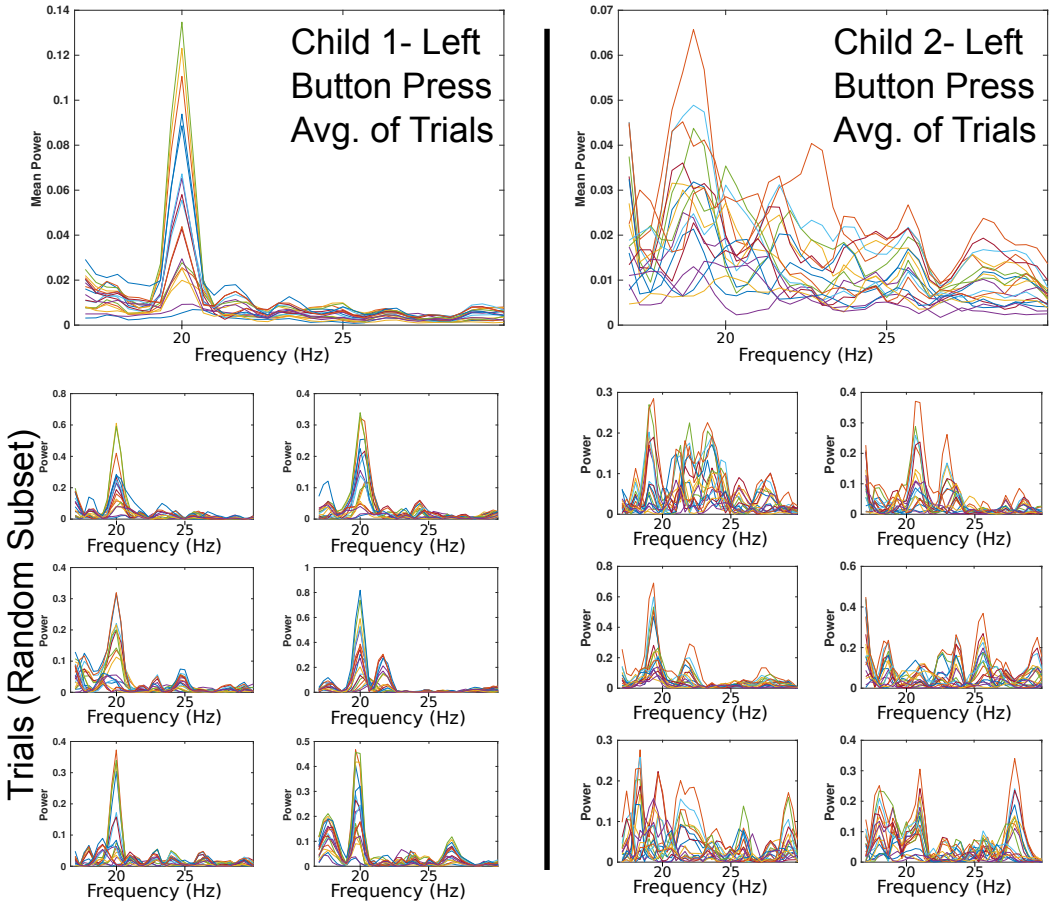

Figure 5.5: Comparison of power at occipital channels between two children during left-selection (20 Hz) tasks in CCP Block 1. The left-selection SSVEP task would expect synchronization at 20 Hz. The left column shows a representative child with a clear left-selection SSVEP signal, while the right column shows a child with less clear left-selection SSVEP signals. The average spectral power across trials for each channel is shown in the top panel of both columns, with the smaller panels illustrating a random subset of the individual trials. Left-task classification accuracy for Child 1 was 100% and for Child 2 was 64% using the proposed tensor analysis.

Also of interest is that significant correlation exists between a child's contribution in a given decomposition and their reported classification accuracy score (Pearson's correlation, with p-values $p = 0.0055, p = 0.0002$ for total accuracy of left and right trials, respectively). However, this correlation held only when considering the signals of interest. No correlation was found between contribution to noise and any accuracy (with p-value, $p = 0.9895$). Correlation was calculated only for the subset of children who were common across all CCP Blocks.

5.5 Discussion

The results from this study outline a tensor analysis scheme for binary classification which is capable of analyzing SSVEP data from children using the direct projection scheme described in Chapter 2. Through projection, the need for a subject-specific calibration step is removed. This reduces some of the set-up and calibration down time associated with traditional BCI tasks, and would thereby improve the user-experience for children, tackling one of the obstacles discussed in Chapter 3. Additionally, through isolating unwanted background noise into a single component, the tensor analysis is able to minimize the impact of continually developing electrophysiological states present throughout childhood [119], tackling another barrier discussed in Chapter 3, and building upon our contributions in Chapter 4. Finally, by including a $[Trial]$ domain, tensor analysis has huge potential for paediatric use because it can emphasize the trials in which a child does well, while minimizing the overall impact of failed, aborted or unsuccessful trials. This has strong implications in BCI paradigms, as children are more likely to abort or fail a task than adults [31, 115, 135].

The tensor analysis showed significant classification capabilities despite potential differences in development between the children. Of note, by ordering the $[Subject]$ domain to be increasing in age from 8-11 years old, general developmental information can be elucidated. For example, it appears that the latter half of subjects in Block 1 (seen in bottom right panel of Figure 5.3) correspond to the largest contributions in each component. For Block 1, subjects 1-17 were age 8-9 years old, while subjects 18-34 were age 10-11 years old. In this regard, contributions to components from the older children trended towards being more dominant, especially regarding the noise factor component, than younger children. From a developmental point-of-view, this indicates that the individual children who performed best in the task (i.e.

contributed most significantly to the CP component model) were, unsurprisingly, often the older children represented by the latter half of the $[Subject]$ domain in Figure 5.3. This is especially obvious for Child 24, Child 28 and Child 31 in Figure 5.3 for left-selection task contribution (blue factor), who all had exceptionally clear left-selection SSVEP.

However, not all of the older children contributed more than the younger subjects. This highlights a few important developmental considerations in the results. First, the child's chronological age may not necessarily reflect their developmental state in its entirety. This concept is explored further in Chapter 6, where additional markers for development are considered. Second, by emphasizing children (and trials) which contributed most to target signals of interest (e.g. the 20 and 25 Hz SSVEP signals) as well as identifying which children were mainly dominated by non-target signals (e.g. noise), the tensor analysis scheme can capture latent information on the developmental state of the children. To this end we postulate that the children ahead (behind) in development are reflected by larger (smaller) contributions to the tensor components. Then, if we consider the strong correlation between relative component contribution in the CP model and the classification accuracy, tracking the development of a child across sessions/interventions/therapies may be possible using the classification accuracy as a guide. This concept could be explored further in future investigations.

Given the simple nature of the classification, i.e. taking whichever new weight w_y is largest, the proposed tensor analysis scheme could be extended easily to decompositions with more than 2 target signals. Additionally, as classification is based on a rapid matrix multiplication of new trials and subjects by the encoding matrix, the proposed work could be readily introduced into real-time BCI systems. The computational complexity required for such endeavours would be relatively low. The major computational bottleneck occurs during the decomposition step, which could be done 'offline', e.g. prior to the 'online' real-time classification element. The specific computational complexity would depend on the algorithm used (see [152] for details), but using tools like Tensorlab provides linear complexity in the CP decomposition algorithm [60]. The online complexity is trivial, as the direct projection required for the getting the new estimated weights for classification only requires simple matrix multiplication. Demonstrating the proposed tensor framework in a real-time scenario of a BCI application is another avenue

for future work.

5.5.1 Limitations

The presented work has some limitations to consider. When designating a specific trial as 'left' or 'right', the recorded button pressed was considered to be the ground truth. However, it is possible that some children may have focused on one object (left or right) and accidentally pressed the other button. As the analysis is based on a public dataset, and minimal direct oversight was given for each trial analyzed, it is possible some trials could have suffered from this discrepancy. However, the effect of such an error is likely to be small considering that the decomposition results provide flexibility in the $[Trial]$ domain for contributions to left/right components. With the significant classification results reported, the overall tensor analysis scheme remains quite promising.

Additionally, the public CMI dataset used for analysis in this chapter was not necessarily developed with the intent of SSVEP analysis and classification. However, the CCP task paradigm was designed in such a way that the relevant information was available for us to pose the questions and analysis presented in this chapter. Subsequent investigations could expand upon this preliminary work through demonstrating its validity in an actual SSVEP-BCI setting, like those seen in [135].

5.6 Conclusion

This chapter outlined a tensor analysis structure for classifying SSVEP signals of interest from a paediatric EEG dataset, expanding upon the findings from Chapter 4. The presented work demonstrated how a 4-mode tensor with domains $[Trial] \times [Spatial] \times [Spectral] \times [Subject]$, could be utilized for analyzing more complex EEG data, e.g. SSVEP-driven EEG. The 4-mode tensor was used to derive a direct classification between two SSVEP trial categories, 'left' and 'right' based on the CCP paradigm. The 4-D tensor was modelled using a constrained 3-component CP decomposition, where orthogonality constraints guaranteed independence of the SSVEP signals of interest. Extracted components from the decomposition were used to define an encoding matrix for projecting new subject and new trial data onto the latent structural interactions found through the CP tensor decomposition. Results found left and right

trial classification was approximately 70% accurate across combinations of training and testing blocks. Additionally, the tensor analysis was shown to be suitable even in children with noisy trials and unclear SSVEP signals as seen in Figure 5.5.

Exploiting the properties inherent in tensor analysis thus may provide a beneficial framework with boons for paediatric focused SSVEP analysis, e.g. in SSVEP-based BCI. Advantages via the outlined tensor framework includes the potential for dealing with varying noise across a developmental range of children, offering a mechanism to account for varying states of development across children. Additionally, the tensor scheme offers a means for rapid classification without a subject-specific calibration phase, improving suitability for paediatric subjects. Together, these results demonstrate the proposed tensor framework as a promising tool for applications using paediatric SSVEP analysis.

Chapter 6

Data Fusion Towards the Joint EEG Development Inference (JEDI) Model

This chapter is built upon the work described in the accepted journal manuscript to *IEEE Transactions on Neural Systems & Rehabilitation Engineering (TNSRE)* [5].

The key contributions of this chapter are:

- Outline and describe the Joint EEG Developmental Inference (JEDI) model structure.
- Apply the JEDI model using a common EEG task (left/right button-press) and relevant clinical cognitive tests of WASI-II, WIAT-IIA and KKQ.
- Demonstrate how the JEDI model can exploit shared information in order to predict developmental scores for new and unseen children based on EEG recordings in a nested cross-validation scheme.

6.1 Introduction

Accounting for the developmental changes in children is vital in adapting neurorehabilitation technologies like BCI [67, 74] to paediatric populations. In Chapter 4, we proposed methodology for tackling some of the hurdles facing translation of BCI to developing children through using multi-way analysis to exploit the natural higher-order structure of paediatric rEEG data in order to identify developmental profiles. In Chapter 5 we extended the multi-way framework to a more active EEG task and introduced an additional mode to the data, the $[Trial]$ domain, in order to capture information in SSVEP-driven EEG tasks for children. Both Chapter 4 and Chapter 5 demonstrated that the tensor framework was capable of

informing on aspects of child development, through identification of unique feature profiles identified and linked across children of different chronological ages. The developed tensor framework provided a foundation for identifying features with respect to child development as a function of age, but did not take into account other potential measures of development, e.g. traditionally used clinical measurements. Acquiring the well-established clinical tests and questionnaires needed to determine a child's development, however, can be resource and time intensive. With many data-driven rehabilitation approaches relying on EEG data, a means to rapidly assess and infer developmental status of children directly from EEG recordings could be critical. Therefore, this chapter aims to expand our tensor framework and models to include information from such clinical tests via data fusion, with a goal to establish tools to infer child development directly from the EEG.

Data fusion offers a means to integrate and analyze data from multiple sources of information through joint or coupled factorization [153–157]. The coupled (i.e. fused) data can provide insights which are not readily apparent from a single source of data [153–157]. This coupling of information has been referred to as multi-view or multi-relational [153], and arises when the features in an observation consist of two or more disjoint sets or views [153]. Importantly, tensor decompositions, like the CP model, can be understood as a special case of joint factorization (i.e. data fusion) where the sets of data are homogeneous [153]. That is, the data sets are equal in size and share at least one common factor [153]. Structured data fusion (SDF) [60, 153] can offer an application agnostic, customizable data fusion framework capable of jointly decomposing various data types, including tensors and matrices. Therefore, data fusion via SDF offers a means compatible within the tensor framework to incorporate in clinical metrics of child development.

EEG data and traditional developmental scores measured through psychometric tests can intuitively be recognized as multi-modal (i.e. multi-relational) data, where the two disjoint sets reveal complimentary information regarding child development. As such, their joint tensor-matrix factorization and SDF [60, 153] could provide novel insights into the underlying developmental information in paediatric EEG, and provide additional boons in terms of inferring development status of children. Therefore, we have termed the SDF of these data types the Joint EEG Development Inference (JEDI) model. To the authors' knowledge, no

such modelling framework has been previously described in data fusion literature. Effective modelling of the multi-relational data associated with child development supports the proposed JEDI model as a potentially useful tool for characterizing developmental status in emerging rehabilitation technologies. This chapter outlines the pillars of the JEDI model, and validates its potential use in predicting developmental scores for children based on a real-world, task-oriented EEG dataset.

6.2 Materials

6.2.1 Datasets

Data for this study was taken from the same publicly available CMI dataset [25] and task paradigm (Contrast Change Paradigm; CCP) as described in Section 5.2.1.

Rather than focusing on the SSVEP-driven task window of the CCP, analysis in this chapter mainly concentrates on the post-decision processing, pre-motor time-frame but includes up to 500 ms after movement is initiated [25]. This slice of the CCP captures the moments spanning from when the subject has decided which object is increasing in contrast, to the pre-motor signals associated with planning an action, and the EEG associated with the actual button press [25].

The post-decision/pre-motor (PD/PM) isolated brain signal was chosen as it could be considered similar in design to the pre-motor imagined movement common in rehabilitation BCI [42]. A small amount of time after the actual button press was also included in order to capture the tail end of the pre-motor signals and their transition into signals associated with real movement. Furthermore, the CCP task required engagement from the child, thus providing brain activity in which developmental differences could be readily investigated [158–160]. Again, the random set of 24 trials (12 left, 12 right) were used to define a block of CCP tasks, and each child attempted to complete 3 of these CCP blocks. (For full details, please see Section 5.2.1, and [25].)

Children were excluded from analysis in this chapter based on two main factors. First, the experimental set-up required each CCP block 1, 2 and 3 to consist of exactly the same children

across all blocks in order to facilitate comparison, validation and testing. Therefore if any child was unable to complete all 3 CCP blocks they were excluded. Second, all children in the analysis needed to have recorded results for the same developmental tests. Children without the required recorded clinical tests and scores (outlined below) were then also excluded. This resulted in a total of $n = 22$ children for the study.

6.2.2 Developmental Tests

Phenotypic data was recorded alongside the paediatric EEG in the CMI dataset. Several developmental tests were recorded for each subject. To narrow the scope of potential developmental metrics, we *a priori* opted to focus on phenotypic developmental scores associated with IQ (i.e. instead of behaviour, social, etc.). Thus, from the available developmental measurements, we chose three psychological tests to investigate. These were:

1. The Wechsler Abbreviated Scale of Intelligence- Second Edition (WASI-II) [161].
2. The Wechsler Individual Achievement Test- Second Edition-Abbreviated (WIAT-IIA) [162].
3. The Kid-KINDL questionnaire (KKQ) [163].

This set of CMI phenotypic data measurements of child cognition included two independent, external measures of development conducted by an expert (WASI-II, WIAT-IIA), and one self-reflective measure of development (KKQ). Thus, our analysis included the full-scale IQ (FSIQ-4) measurement from WASI-II, the full composite score for WIAT-IIA, and only the School Subscale in the KK questionnaire. The score values for WASI-II and WIAT-IIA were reported on a normative 100 point curve (where 'normal' is considered at $\mu = 100$, $s.d. \pm 10$). The KK questionnaire, however, ranged from 0-5 points for the School Subscale based on responses 'never', 'seldom', 'sometimes', 'often' and 'all of the time', respectively. The School Subscale scores were arranged by the CMI dataset so that higher points indicated the child feeling more successful/positive in the school setting. To more easily visualize across data, the KK questionnaire's 5 point scale was translated to a 0-100 point scale for analysis. Through these three measures we aimed to capture a well-rounded set of classic clinical developmental metrics of child cognition.

The scaled scores from the WASI-II, WIAT-IIA and KKQ were aggregated together across all children to construct a psychometric score matrix, $\mathbf{S} \in \mathbb{R}^{N \times L}$. Here, the $[Subject]$ and $[Score]$ domains are of size N and L, respectively. Prior to cross-validation analysis, the score matrix was standardized to a z-score using Matlab so that the appropriate scale of each of the scores could be recovered in the prediction step.

6.3 Methods

6.3.1 Data Processing

EEG data was originally recorded and pre-processed by the CMI as described in Chapter 5 [25]. Like in Chapter 5, we opted to use the pre-processed EEG data from the CMI to help promote reproducibility of the analysis. For this chapter, the desired EEG recordings from the CCP were re-processed using the Fieldtrip toolbox [58] in Matlab to find the time-frequency response (TFR). First, the data was re-referenced to a common average reference, and bandpass filtered between 5 to 18 Hz. This frequency range was selected for two purposes. First, it captures the variable range of EEG frequencies in children associated with executing motor movement in the button-press task, e.g. the shifts in frequency and power of EEG bands alpha, mu and beta which are associated with development throughout childhood [21, 32, 119, 164] as discussed in Section 2.3.1. The lower beta range is included as it continues to change and mature with the development of the child [32, 119, 164, 165] and is associated with motor activity [32, 164, 165]. The upper beta range was excluded, however, to avoid potential contamination from the steady-state visual evoked signal at 20 and 25 Hz [25] present in the CCP task.

A spatial filter was applied to focus on 20 central channels over the motor cortex (E103-106, E110-112, E116-118 for the right, and E7, E13, E20, E28-30, E34-36, E41 for the left). The TFR was then calculated using Fieldtrip's multitaper convolution function ('mtmconvol'), across Hanning tapers on the time series trial data [58]. The multitaper convolution function convolves data in the EEG time series using a complex wavelet. In particular, the wavelet is constructed by time-point wise multiplication of the (real) cosine and (imaginary) sine component at each frequency within the specific tapering function, e.g. the Hanning taper in this case. For more details, please refer to Fieldtrip's documentation presented in Section

2.8.1. The data was time-locked to the button press response, designated as time $t = 0$. Sliding time windows of 20 ms were used to estimate the TFR for two second trials defined from -1500 to 500 ms, with respect to the button press. A baseline from -1500 to -1000 ms was used to calculate the relative change in the TFR spectrum data. Analysis was done separately for left/right button press data from the CCP.

6.3.2 Tensor Analysis

Figure 6.1-A illustrates a general example of a three-way tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times N}$ corresponding to $[Spatial]$, $[Spectral]$ and $[Subject]$ domains, respectively. This figure is included to facilitate comparison to a 5-way tensor (Figure 6.1-B) and data fusion illustration (Figure 6.1-C).

Constructing the Multi-way Tensor

From the processed EEG data, a 5-way tensor was constructed with domains of $[Trial] \times [Spatial] \times [Spectral] \times [Temporal] \times [Subject]$ using the TFR power spectrum data calculated for each child, at each trial across a CCP block. This 5-way tensor is visually represented in Figure 6.1-B. A subset of the constructed 5-way tensor was identified to be core in defining the 'tensor-aspect' of the joint model. In particular, the tensor model did not need resolution at the $[Trial]$ domain level, as the child's developmental scores did not change from trial to trial. Furthermore, training and validation was done using characteristic filters and direct projection similar in practice to Section 5.3.3 on combinations of CCP recording blocks (e.g. training on CCP block 1 and 2, CCP block 1 and 3, etc.). Thus, a sub-tensor which averaged across the whole $[Trial]$ domain was defined for experimental analysis. Therefore, analysis was conducted on the 4-way sub-tensor $\mathcal{X}_{sub} \in \mathbb{R}^{I \times J \times K \times N}$ corresponding to $[Spatial] \times [Spectral] \times [Temporal] \times [Subject]$ domains, respectively, for each CCP block.

The constructed 4-way tensor was decomposed via the Tensorlab toolbox [60, 166, 167] for Matlab, using the built in 'CPD' (i.e. CP model) function. The CP model was again chosen for similar reasons stated in Chapter 2, 4 and 5, e.g. its mild requirement for uniqueness, the ability to exploit the guaranteed 1:1 interaction for identified latent factors across the domains of $\mathcal{X}$, to

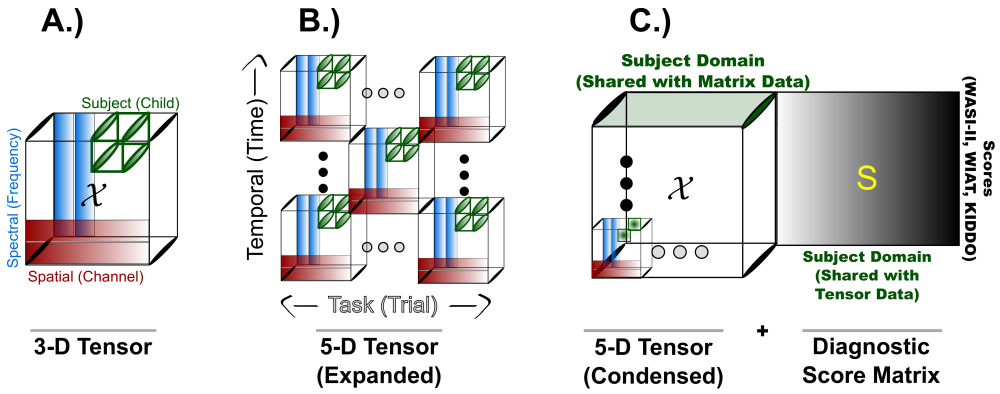

Figure 6.1: Visual representation of various tensor structures. A.) A simple representation of a 3-domain tensor $\mathcal{X}$ with $[Spatial]$ (channel), $[Spectral]$ (frequency) and $[Subject]$ (child) domains, included for convenience. B.) Illustration of how the 3-way tensor from (A) can be stacked along $[Temporal]$ (filled dot) or $[Task]$ (open dot) modes to create a 5-domain tensor. C.) An illustration of the shared $[Subject]$ domain between a 5-way tensor $\mathcal{X}$ and a matrix $\boldsymbol{S}$ (e.g. developmental score matrix).

utilize the characteristic filters, etc. Using the smaller 4-way sub-tensor in the analysis provided an additional benefit in the form of reducing computational expense when decomposing the tensor facet of the tensor-matrix data fusion.

6.3.3 Data Fusion and Grid Search

The EEG tensor and the score matrix of were jointly decomposed using the CPD option (i.e. CP model) in Tensorlab's structured data fusion (SDF) [60, 153]. Explicitly, the tensor and matrix data arrays had the forms of:

$$\mathcal{X}_{sub} = [Spatial] \times [Spectral] \times [Temporal] \times [Subject] \in \mathbb{R}^{I \times J \times K \times N} \tag{6.1}$$

$$\mathbf{S} = [Subject] \times [Score] \in \mathbb{R}^{N \times L} \tag{6.2}$$

The component factors for the $[Subject]$ and $[Score]$ modes in the score matrix $\mathbf{S}$ were $\mathbf{D} \in \mathbb{R}^{N \times R}$ and $\mathbf{E} \in \mathbb{R}^{L \times R}$, respectively. The score matrix $\mathbf{S}$ also had the property:

$$\mathbf{S} \approx \mathbf{D}\mathbf{E}^T \tag{6.3}$$

The shared $[Subject] \in \mathbb{R}^N$ domain, seen in equations (6.1) and (6.2), was the single overlapping domain from the multi-relational set given to the SDF model. Figure 6.1-C visually illustrates the shared domain in a tensor-matrix SDF. An initial estimate for the decomposed tensor was first calculated using Tensorlab's multi-linear SVD [166] while the score matrix was initialized randomly.

Non-negativity was imposed on each domain of the EEG tensor (and thus consequently on the $[Subject]$ domain of the score matrix). Similar to discussion Chapter 4 and 5, the non-negativity constraint improved the ability to interpret factors of the model, grounding them within realistic and interpretable bounds. Regularization was imposed on both the EEG tensor and score matrix data. For the EEG tensor, L2 regularization helped reduce fitting to noise for the $[Spatial]$, $[Spectral]$, and $[Temporal]$ domains. Similarly, L2 regularization was included for the matrix's $[Score]$ domain. L1 regularization was imposed on the $[Subject]$ domain for both datasets to help promote minimal over-fitting and potentially sparse responses across the shared mode.

Using the SDF structure, the degree of contribution of the EEG tensor, score matrix, L2 and L1 regularization terms were defined in terms of their relative weights. These weights can be considered as hyperparameters of the JEDI model. The objective function to minimize the SDF of $\mathcal{X}$ and $\mathbf{S}$ is thus:

$$\begin{aligned} \min_{\mathbf{A,\,B,\,C,\,D,\,E},R} \quad & (\lambda_1/2)||\mathcal{X} - \mathcal{M}_{CP}(\mathbf{A,\,B,\,C,\,D}, R)||_F^2 + \\ & (\lambda_2/2)||\mathbf{S} - \mathcal{M}_{CP}(\mathbf{D,\,E}, R)||_F^2 + \\ & (\lambda_3/2)(||vec(\mathbf{A})||_F^2 + ||vec(\mathbf{B})||_F^2 \\ & + ||vec(\mathbf{C})||_F^2 + ||vec(\mathbf{E})||_F^2) + \\ & (\lambda_4/2)||vec(\mathbf{D})||_1 \end{aligned} \tag{6.4}$$

where **A, B, C, D, E**, R are the $[Spatial]$, $[Spectral]$, $[Temporal]$, $[Subject]$, and $[Score]$ domain component matrices and number of components, respectively; $\mathcal{M}_{CP}$ is the joint CP model decomposition; and λ_{1-4} are the relative weights for tensor data $\mathcal{X}$, the score matrix $\mathbf{S}$,

L2, and L1 regularization, respectively.

In order to estimate λ_{1-4} in (6.4), we employed a grid search varying λ_{1-3} while fixing λ_4 at 1. We varied $\lambda_{1,3}$ logarithmically from 0.01 to 100 in nine steps each (i.e. 0.01, 0.0316, 0.1, 0.316...). To be sure that the relative weight of neither the tensor nor the matrix was dominant, we varied λ_2 logarithmically proportional to λ_1 such that it was at most (least) one log-step more (less) than the value for λ_1 (i.e., for $\lambda_1 = 1$, $\lambda_2 = [0.1, 10]$), over 5 steps in log-space. For each value λ_{1-4}, the SDF was calculated for one of the three CCP blocks (e.g. CCP block 1), and validated on an independent second CCP block (e.g. CCP block 2). Validation using the second CCP block (without loss of generality) provided the relative errors for each combination of parameters. This procedure thereby approximated the generalizability of each set of hyperparameters with respect to unseen EEG data, providing insight for choosing an ideal set of hyperparameters applicable across all combinations of the CCP blocks. The third CCP block not used for either defining the initial SDF model or grid search validation was used for testing (e.g. CCP block 3).

In addition to the relative weights λ_{1-4}, the grid search varied the number of components *R* for the joint CPD decomposition between 2 to 5 components. This range for the components was chosen specifically to emphasize the tensor ranks in which the JEDI model effectively balances (1) model validity measured by percent explained variance (EV); (2) model validity of the CPD using CORCONDIA, similar to Chapter 4; and (3) computational time required. As discussed in Chapter 4 CORCONDIA indicates the suitability of using CPD to model the data and is inversely related to the EV [143, 168]. A larger number of components used for a CPD decomposition tends to decrease the CORCONDIA while increasing the EV and the computational time [52, 167]. As such, selecting the correct number of components to decompose a dataset requires a balance between these aspects.

The grid search was ran over all permutations of the independent CCP blocks 1, 2 and 3 for model definition, validation and testing. Results were recorded and a specific model weight with *R* components was selected for analysis.

6.3.4 Extending the Characteristic Filters and Direct Projection

In this study, we again exploited the CP model structure to utilize characteristic filters and direct projection techniques, outlined in Sections 2.6, 4.3.5 and 5.3.3. Here, we modified the characteristic filters to include the $[Spatial]$, $[Spectral]$, and $[Temporal]$ domains in order to estimate weights for new children projected onto the $[Subject]$ domain. Using the common interactions held by the characteristic filters of the $[Spatial]$, $[Spectral]$, and $[Temporal]$ modes in $\mathcal{X}_{sub}$, we can estimate the associated component weights of new children in the $[Subject]$ domain directly, e.g. for the unseen testing CCP-block.

The characteristic filters developed in this chapter expanded upon those previously described, as an additional mode was included in the direct projection (i.e. the $[Temporal]$ domain). Details on this variation are presented here for completeness.

Let $\mathcal{X}_{sub} \in \mathbb{R}^{I \times J \times K \times N}$ and $\mathbf{S} \in \mathbb{R}^{N \times L}$ again be the 4-way sub-tensor and 2-way score matrix, respectively, with resulting *R*-component factor matrices as described in Section 6.3.3. Then, let a new testing tensor $\mathcal{Y}$ be of similar design to $\mathcal{X}$ (i.e. have the same domain designations), with new subjects and recordings. Then weights for the new subjects in the $[Subject]$ domain of $\mathcal{Y}$ (i.e. $\hat{\mathbf{D}}$) can be determined through direct projection on the characteristic filters of $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ as follows:

- First, unfold $\mathcal{X}_{sub}$ along the $[Subject]$ domain, defining the $[Spatial] \times [Spectral] \times [Temporal]$ characteristic filters using the Khatri-Rao product:

$$\mathbf{X}_{sub}^{(4)} = \mathbf{D}[\mathbf{C} \odot \mathbf{B} \odot \mathbf{A}]^T, \tag{6.5}$$

$$\mathbf{X}_{sub}^{(4)} = \mathbf{DW} \tag{6.6}$$

  where $\mathbf{W}$ is the encoding matrix of the interactions between the $[Spatial]$, $[Spectral]$, and $[Temporal]$ domains [50]. The projection from new subjects onto these common interactions provides estimated weights for the subjects' spatial, spectral and temporal profile patterns [50].

- New weights from the direct projection are then found using the non-negative

least-square projection ($NN(\cdot)$) between the encoding matrix, $\mathbf{W}$, and the unfolded testing tensor $\mathcal{Y}$ along the $[Subject]$ domain, $\mathbf{Y}^{(4)}$, similar to Section 2.6:

$$\mathbf{Y}^{(4)} = \mathbf{D}'[\mathbf{C}' \odot \mathbf{B}' \odot \mathbf{A}']^T, \tag{6.7}$$

$$\hat{\mathbf{D}} = \text{NN}(\mathbf{Y}^{(4)}\mathbf{W}), \tag{6.8}$$

where $\hat{\mathbf{D}}$ holds the estimated channel, frequency, and time weights for the new subjects [50]. The non-negative least squares as an alternative to the pseudo-inverse to inhibit potential outliers from near-zero data, similar to Chapter 4.

- Since $\mathcal{X}_{sub}$ is jointly factorized with the developmental score matrix $\mathbf{S}$, the new estimated $\hat{\mathbf{D}}$ will be encoded with respect to the developmental scores as well. Extending the characteristic filter outline presented in Section 2.6 further, we can use the direct projection technique to estimate a new child's estimated developmental score for WASI-II, WIAT-IIA or KKQ. We need to simply multiply the estimated weight $\hat{\mathbf{D}} \in \mathbb{R}^{N \times R}$ by the transpose of the component matrix found for the $[Score]$ domain, $\mathbf{E} \in \mathbb{R}^{L \times R}$.

$$\hat{\mathbf{S}} = \hat{\mathbf{D}} \times [\mathbf{E}]^T \tag{6.9}$$

where $\hat{\mathbf{S}} \in \mathbb{R}^{N \times L}$ is the new estimated score for the child(ren).

6.3.5 Cross-Validation Sets

The above analysis was done on a combination of the 3 CCP blocks, such that analysis included evaluation on a cross-validated test using two CCP blocks and a repeated evaluation on the other withheld CCP block. In particular, a 10-fold cross-validation combining two CCP blocks of data withheld two or three subjects at each fold when training the JEDI model to establish the "within block" validation set. All data from the other CCP block was completely excluded from the model training as well, to provide the second validation test set designated as the "out-of-block" test. This double hold-out scheme promoted robustness in the results two ways. First, evaluating two repeated tests helps inform on the stability of the JEDI model across multiple evaluations. Second, by excluding a CCP block from the model training, potential

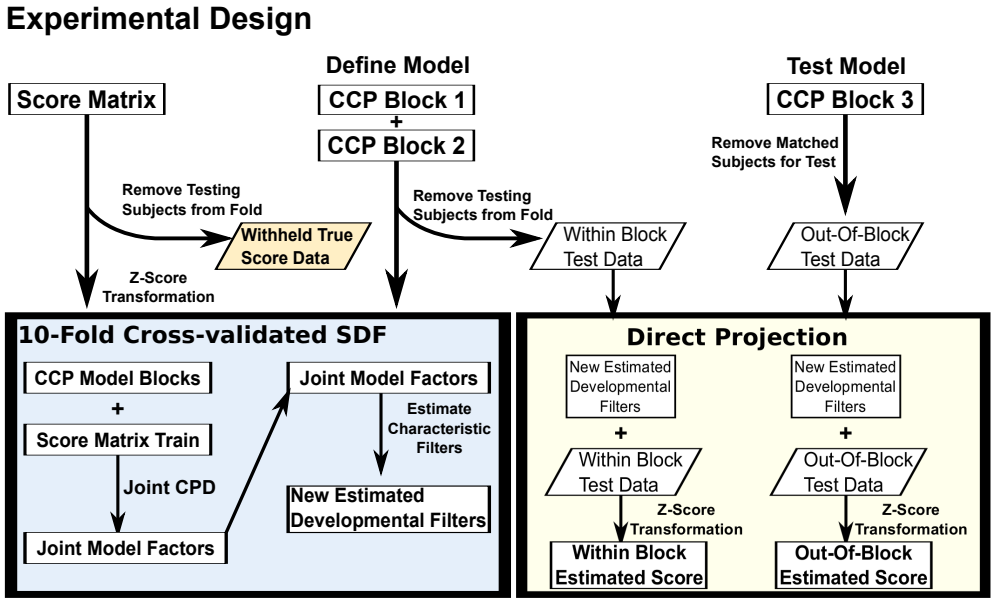

Figure 6.2: A flowchart illustrating the experimental design of the JEDI training and testing. The flowchart is read from top to bottom, left to right. First, a Score Matrix and set of CCP Blocks are used to build the tensor model. They are processed as shown, through the arrows and blue box. At the same time, withheld data from the 3rd CCP block is processed so it can be used for the testing section. The characteristic filters derived from the blue box are used in the direct projection shown in the yellow box. Importantly, the true scores for the data are withheld prior to data processing, as highlighted in the tan parallelogram. Overall, the blue box highlights the key steps in the JEDI joint factorization for each training fold. The yellow box highlights how withheld testing data was used to estimate scores for both the 10-fold cross-validation and the out-of-block independent test data.

systemic differences present in the CCP blocks may be revealed from comparing the within block and out-of-block validation results. Importantly, the children to be tested had their scores held out prior to the z-score transformation as well, to ensure no potential contamination between the training and testing sets. Figure 6.2 illustrates this experimental outline. The new estimated scores for both within the cross-validation and out-of-block testing were compared to the original true scores of each child. The joint CPD factorization was run on each training fold separately, and repeated using all permutations of the CCP blocks 1, 2, and 3.

6.4 Results

6.4.1 Parameter Weights

The joint factorization reported for the JEDI model used $R = 4$ components, and hyperparameter weights of 3.162, 3.162, 0.1, and 1 for λ_{1-4}, respectively. These values were chosen based in part on results from the complete grid search. To maintain commonality across the various testing experiments, we opted to use a single set of weights, selected from the set of parameters which gave the five smallest differences between real and estimated values for the decomposition and validation blocks in the grid search. As an example, we would select the weights from parameters minimizing differences between the real and estimated values based on using CCP block 1 for joint decomposition and CCP block 2 for validation, while keeping CCP block 3 unseen for use in the independent testing. All permutations of this order were investigated, with the set of chosen hyperparameter weights selected from the 5 smallest differences from all permutations. Of these hyperparameters, those using 2 component factors were not considered for the analysis, as the limited number of components were unable to effectively capture expected trends in the data based on visual inspection. Similarly, hyperparameter values near an extreme end of the grid search (e.g. $\lambda_1 = 10$, $\lambda_2 = 100$, $\lambda_3 = 31$, etc.) were investigated, but ultimately excluded for use, as they resulted in massive errors in the JEDI model's ability to accurately predict scores. This was likely a result of the extreme values effectively ignoring the tensor, matrix, L2 or L1 regularization aspects of the SDF data within the joint factorization.

Using the single set of hyperparameter weights allowed direct comparisons of results across

permutations, scores and testing blocks since each JEDI joint factorization weighted the profiles identically.

6.4.2 Developmental Score Prediction

Figure 6.3 illustrates the results for the predicted developmental scores, based on the permutations of model training and testing from CCP blocks 1, 2, and 3 using the procedure outlined in equations (6.5)-(6.9). The three permutations of the JEDI model training and testing are represented in each row of Figure 6.3. The left hand column illustrates the predicted scores for using EEG from the 'within block' of the cross-validated data as outlined in Figure 6.2. The right column shows the predicted score results using EEG data from the independent 'out-of-block' testing fold, as seen in 6.2. For each panel, the actual and predicted scores for each child is displayed side-by-side for the WASI-II (left-most data), WIAT-IIA (central data), and KKQ (right-most data) clinical tests. Box-plots were used to capture overall trends in the real and estimated scores. A scatter-plot linking real-to-predicted data was overlaid on top the associated box-plot pairs in order to show the efficacy of the JEDI model at an individual level.

The majority of trends in Figure 6.3 indicate the JEDI could estimate a child's developmental score to a reasonable degree (see Section 6.5.1). For the WASI-II test, a median difference of ± 20.81, and ± 21.68 points on average across permutations for within and out-of-block tests, respectively. Similarly, for the WIAT-IIA test a median difference of ± 15.99, and ± 16.85 points across the different permutations for the within and out-of-block tests, respectively. The KK questionnaire had the lowest median difference in points at ± 8.89, and ± 11.07 across permutations for the within and out-of-block tests, respectively. The lowest median value for the KK questionnaire was ± 4.79 (in the within block cross-validation test), while the highest was ± 13.91 in the out-of-block test. These results are summarized in Figure 6.4 as a complementary graphic to the main results in Figure 6.3.

Across all model testing permutations, only 1 pair of real and estimated scores were found to be significantly different based on a two-tailed Student's *t*-test, at a level of $p = 0.05$. This was the WIAT-IIA scores for the within block at $p \approx 0.0180$, using CCP blocks 1, 3 for training and block 2 for testing. However, correcting for multiple comparisons via False Discovery Rate (FDR) or Bonferroni found no significant differences from the distribution of real scores

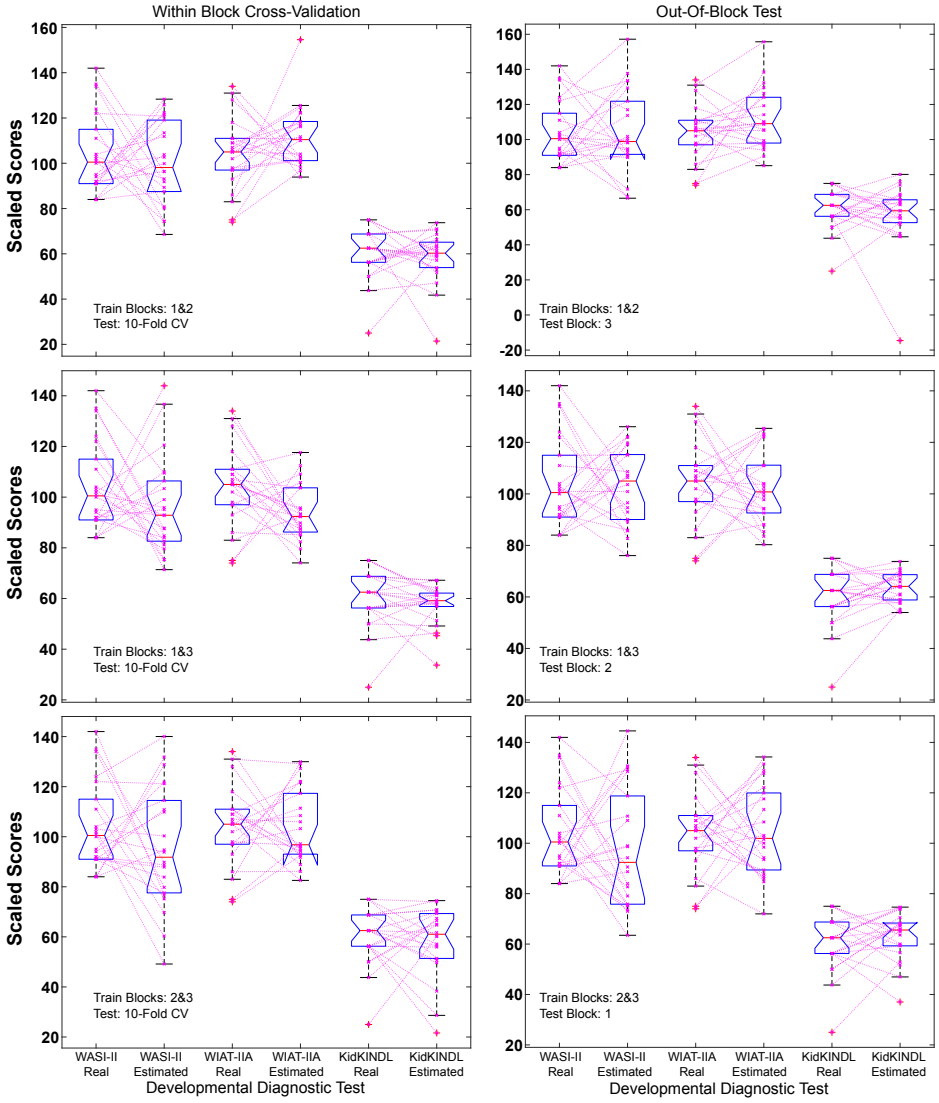

Figure 6.3: Actual and predicted scores for all subjects for three clinical measures of development, for each training and testing permutation of CCP blocks for the JEDI model. Rows reflect the training/testing permutation. The left column gives results from the 10-fold within block cross-validation, and the right column for the out-of-block data. For each panel, from left to right are actual and predicted value pairs for WASI-II, WIAT-IIA and the KidKINDL questionnaire, respectively. Each pair of box-plots includes the subject distribution in magenta, with a dotted line indicating the child's 'actual-to-estimated' score values. No significant differences between the actual and predicted values were found, based on a two-tailed Student's t-test at a significance level of $p = 0.05$*, corrected for multiple comparisons via False Discovery Rate (FDR, at* $p = 0.05$*)*

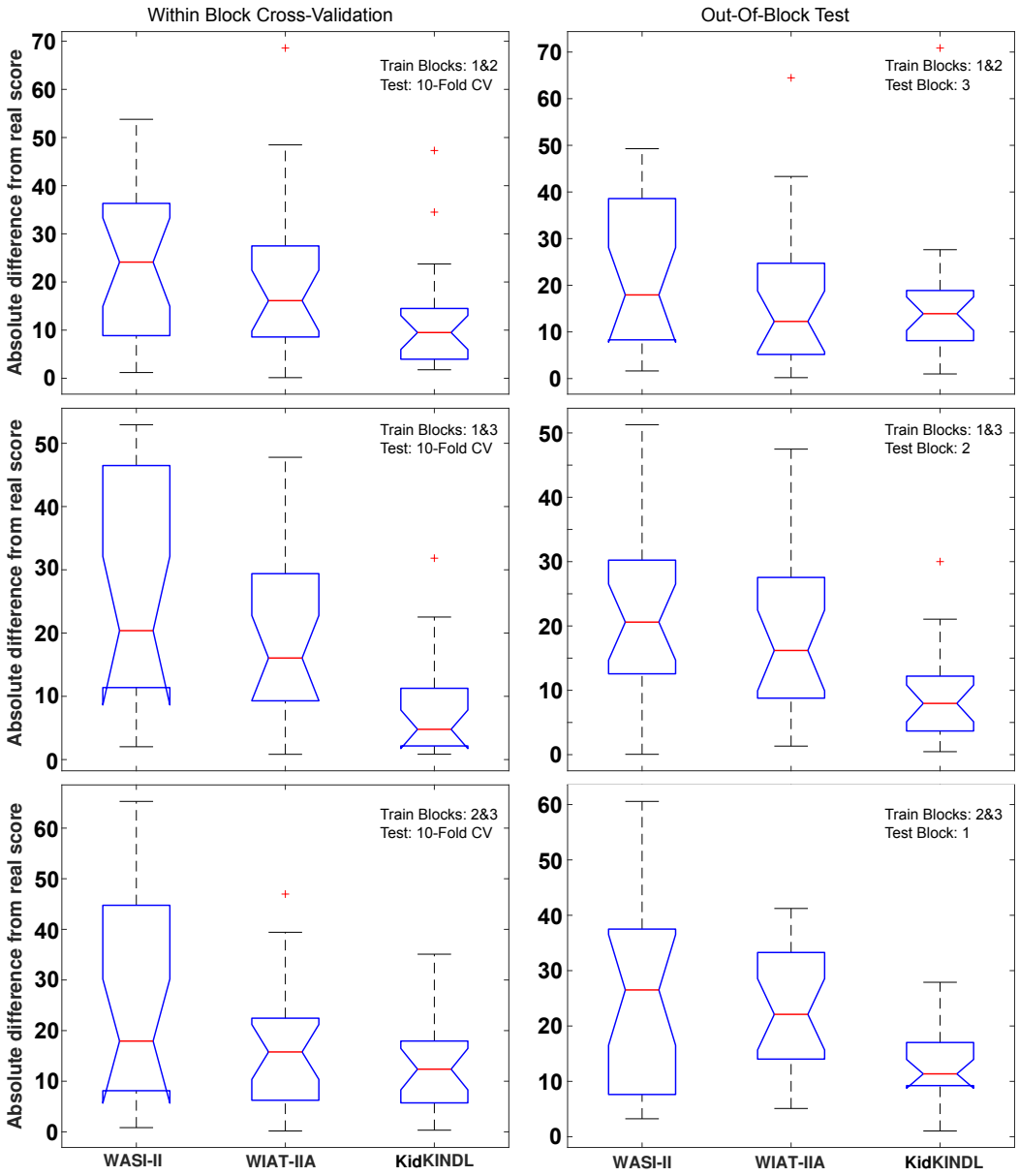

Figure 6.4: Absolute difference between the real and estimated scores given by the JEDI model, for all permutations of CCP blocks. The three left-hand panels show results using the within block cross-validation, while the right-hand panel shows results using the out-of-block test data. For each panel, from left to right the developmental tests are: WASI-II, WIAT-IIA and KKQ (KidKINDL).

to the JEDI estimated values.

Incorporating a child's age into the score matrix **S** was also examined as a measure of verifying the suitability of the JEDI model in estimating aspects of child development. Including child age as an alternative score in the matrix **S** was found to have no significant effect on score estimations made by the JEDI model. Further, the median estimated ages differed from the real ages by ± 1.36 and ± 1.12 on average across all permutations for the within block and out-of-block tests, respectively.

6.4.3 Profiles from the JEDI Model

To demonstrate the tensor aspect of the JEDI model still characterizes important properties underlying the paediatric EEG, an example set of factor profiles from the joint factorization are presented in compactly in Figure 6.5. This example is taken from training and validation using CCP blocks 1 and 3 of the right button press data respectively, leaving CCP block 2 right button press for unseen data testing. The identified component matrices from the tensor $\mathcal{X}_{sub}$ in the joint factorization are illustrated in Figure 6.5-A. The associated factors profiles resemble those seen in Chapters 4 and 5, with the $[Spatial]$, $[Spectral]$ and now $[Temporal]$ modes useful for interpreting what phenomena in the EEG was effectively captured by the tensor decomposition. Channels 1-10 in the $[Spatial]$ domain correspond to the central left brain hemisphere, with channels 11-20 corresponding to the central right brain hemisphere. Time in the $[Temporal]$ mode is set relative to when the button was pressed (i.e. $t = 0$ seconds).

Additionally, the $[Score]$ factors resulting from the joint decomposition in the JEDI model are included in the compact Figure 6.5-B to highlight how the 1:1 interaction in the CP model is carried over to the score matrix, **S**. The shared $[Subject]$ domain of the JEDI model is repeated for emphasis, to show it is the linking mode between the two sets of data. Similar profiles can be identified for each combination of CCP blocks, for both the right and left button press.

An expanded version of Figure 6.5 is given in 6.6, for ease of comparison to profiles presented in Chapters 4 and 5.

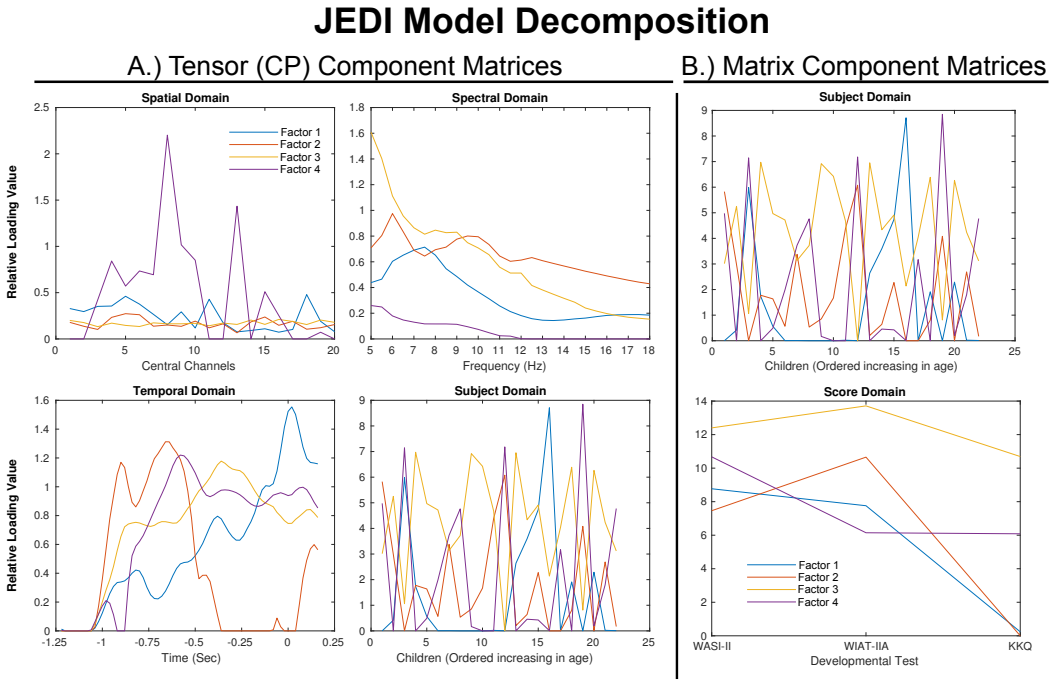

Figure 6.5: An example of the factor profiles from the resulting component matrices using the JEDI model joint decomposition. The profiles given represent data taken from training and validation on CCP blocks 1 and 3, with testing to be done on CCP block 2. A.) Resulting component matrices from the tensor side of the joint factorization, $\mathcal{X}_{sub}$. Factor profiles are shown for the $[Spatial]$, $[Spectral]$, $[Temporal]$ and $[Subject]$ modes based on the hyperparameters found by the grid search. B.) The corresponding factor profiles for the $[Subject]$ and $[Score]$ domains from the score matrix $\mathbf{S}$ *from the joint factorization are displayed.*

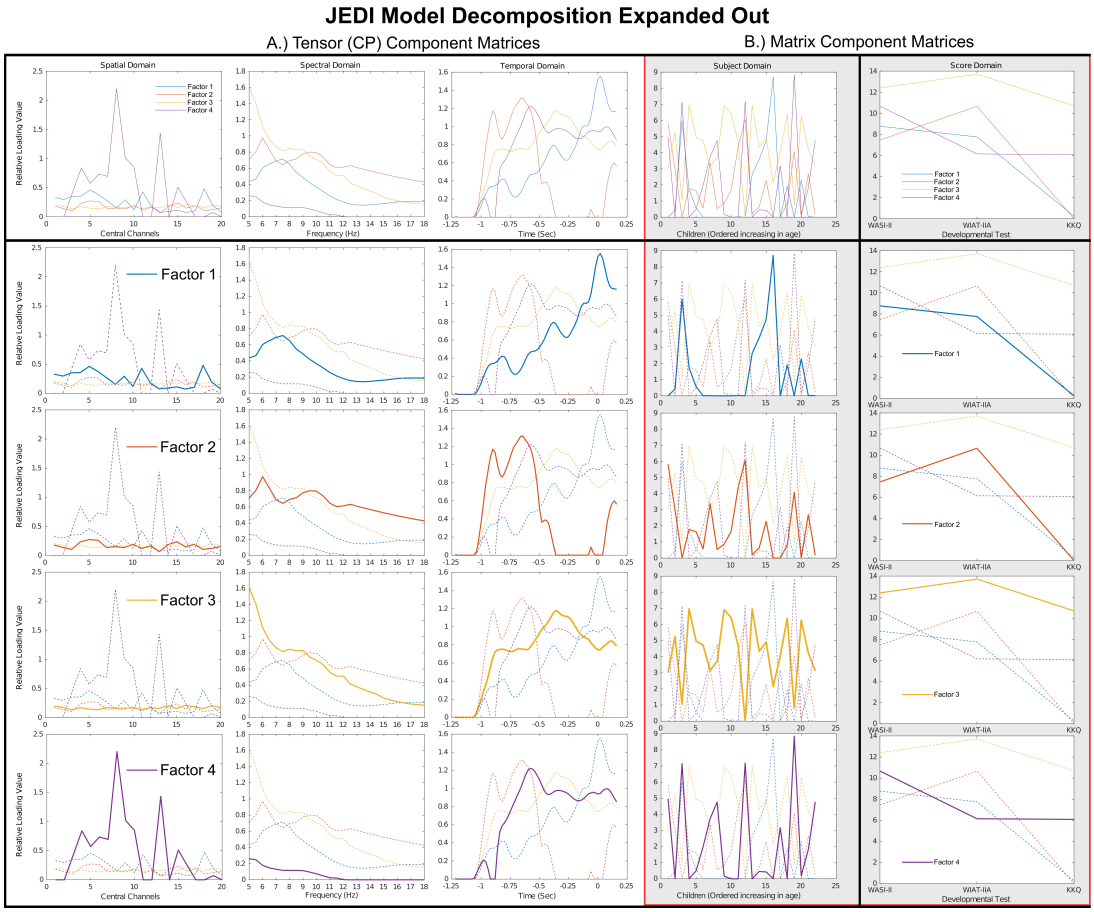

Figure 6.6: An illustration of the expanded profiles resulting from component matrices using the JEDI model joint decomposition. The resulting component matrices are expanded to individually highlight Factors 1-4 for each of the data modes. From left to right, the columns show $[Spatial]$, $[Spectral]$, $[Temporal]$, $[Subject]$ *and* $[Score]$ *modes, respectively. The top most row shows the combined factor matrices, while each row below highlights a particular factor, e.g. Factor 1, Factor 2, etc. The corresponding factor profiles found via the joint decomposition from the score matrix* $\mathbf{S}$ *are highlighted in a red box, in the last two columns of the figure.*

6.5 Discussion

This chapter presents a new model built upon the join factorization of EEG data and clinical measures of development with the aim to better infer developmental status of children using EEG data. The JEDI model was described and validated on predicting WASI-II, WIAT-IIA and KKQ scores rapidly via characteristic filters. The results demonstrated that the JEDI model can estimate a child's developmental score within an adjacent category (e.g. 'average','above average', etc.) of the real data. On a population level, the distribution of developmental scores predicted by the JEDI model were not significantly different from the original distribution of scores in the paediatric population. Additionally, the factor profiles identified from the JEDI model are still capable of capturing key underlying features present in the paediatric EEG data.

6.5.1 JEDI Trends

Results from the JEDI model are an encouraging, initial step. Although the median values across the tests were modest with respect to predicting scores, there were some important trends in the data which promote further investigation. For example, across the WASI-II and WIAT-IIA tests, there appeared to be a trend for individual subjects in the interquartile range (i.e. the 'middle' 50% of data) to have lower differences between the real and estimated scores. In contrast, the more 'extreme' scores (either above or below) tended to be inversely estimated (e.g. low scores were estimated significantly higher by the JEDI model, while high scores were estimated significantly lower). Pearson's *R* reveals a negative correlation trend between the estimated and real scores in the WASI-II (within block testing $R = -0.3403 \pm 0.138$, and out-of-block testing $R = -0.1278 \pm 0.1694$) and WIAT-IIA (within block testing $R = -0.2639 \pm 0.2019$, and out-of-block testing $R = -0.1073 \pm 0.1425$) across all CCP block combinations. These trends are significant for the uncorrected $p - values$, but fail to remain significant when corrected for multiple comparisons for a FDR of $p = 0.05$. This helps explain, in part, why many of the large differences in individual estimates of developmental scores tended to come from data originally in the top or bottom 25% of the population distribution. This trend, however, is not apparent for the KK questionnaire, where estimated data appear to be more positively correlated at $R = 0.1411 \pm 0.2627$, and $R = 0.1805 \pm 0.1038$ for within and out-of-block testing respectively. Furthermore, the KK questionnaire data seem

more likely to be in a similar range as the real data regardless of their positioning in the original data distribution.

It is important to contextualize the estimated results and differences from the JEDI model with respect to a real-world interpretation. In practice, the median differences found between the real and estimated score pairs for WASI-II demonstrate a reasonable ability for the JEDI to retain a child's 'descriptive category' with respect to the original FSIQ-4 values. The median change of approximately ± 21 points for JEDI model estimates of WASI-II imply that children assigned an 'average' score can be categorized as 'below average', 'average', or even 'above average'. The normal curve distribution of WASI-II has a standard deviation of 15 points, putting the JEDI estimation outside the expected deviation. The consistency from both the within block cross-validation and out-of-block test data supports that, given the specific parameters for the presented JEDI model, these results are likely stable.

The WIAT-IIA predicted scores were consistently closer to their real counterparts via the JEDI model. Given that the standard deviation for WIAT-IIA is reported at 15 points, the JEDI model median differences fell critically within only 1-2 points of the expected standard deviation curve, evaluated for the within and out-of-block tests, respectively. This implies the shared developmental factors underlying the EEG were more readily captured via the JEDI joint factorization. This effect seemed to be constant for both within block cross-validation and for out-of-block testing, indicating it also was a stable estimation.

The estimated KidKINDL questionnaire values were also consistently within one categorical step of their respective real counterpart. The within block cross-validation and out-of-block testing median differences seem slightly less stable than the WASI-II or WIAT-IIA predictions (e.g. the approximate 4 point difference in median change). When scaled back to a 5 point scale the predictive capability seems comparable to the other tests, in that estimations will fall within the 1-2 category range (median change of ± 1.36, and ± 1.44 for within and out-of-block testing, respectively). This instability could be partially a result from the specific model joint factorization properties, which could change with different model selection (e.g. different weights of λ_{1-4}, R, etc.).

For all developmental tests, the outlier results represent a key situation where the model

completely failed. From a technical point of view, these failures are likely tied to when components in the model are near zero-valued for an extensive set of elements in any mode, i.e. the zero-values of Factor 2 in the $[Temporal]$ domain of Figure 6.5. This phenomena is more prevalent when the joint factorization does not readily converge to a suitable model, potentially due to local minima or slow convergence swamps [169]. The strong zero presence in the components may then become magnified during the direct projection step, depending on the specific contribution and combination of components in the $[Subject]$ domain, leading to wildly inaccurate scores. Using a more flexible tensor model or improving the optimization of the joint factorization are two potential avenues which may help address this drawback.

However, from a clinical point of view, these outliers are such drastic failures that they are easy to immediately identify. Negative developmental scores (seen in several predicted children in Figure 6.3) are clearly impossible, so a decision can be made to investigate with the classical approach of running the developmental tests. Participants with scores near the middle of the data however could potentially have their development appropriately estimated from the JEDI model, without needing to run the time-intensive clinical tests. The JEDI model therefore could help to reduce the resources required for clinical developmental assessment.

6.5.2 Model Selection and Stability

Model selection is a critical aspect to consider when evaluating the given results. We opted to use a single set of hyperparameters across all combinations of data to retain the ability to more intuitively compare across outcomes. This allowed less complex interpretation when comparing data, at the trade-off of perhaps better fitting of the JEDI model to each training/testing fold and permutation. Selecting appropriate hyperparameters to fit the JEDI model at each fold in turn may improve the estimated scores. This is a potential avenue for additional research, which would require careful consideration of how to best balance the contributing weights of the tensor and matrix data similarly across and within the various data folds, in order to interpret the final results.

Another example of where model selection should be carefully considered, is at the tensor decomposition level. We opted to select the CP model for its ability to retain easy interpretation of features, and exploit the characteristic filters through direct projection as outlined in Section

2.5.3 and 2.6. However, due to the CP model's super-diagonal core [3, 46], the model may be less flexible than desired. Using a more flexible tensor model, like Tucker (described in Section 2.5.4), PARAFAC2 [42, 50] or block-term decomposition [37, 170] offer directions for future work which could improve the estimations from the JEDI model.

Stability in the JEDI model is another key aspect to consider. Despite testing with two separate scenarios, including a completely 'unseen' EEG data, the median results from the JEDI model remained relatively stable. That is, similar results were found for both within and out-of-block testing groups. The general stability of the model is further supported in that none of the estimated distributions were significantly different from the actual values, with distributions found to be approximately similar in size and location as seen in Figure 6.3. Importantly, however, is that the individual predictions can vary if the JEDI model is run repeatedly due to properties in the joint tensor factorization algorithm and the random initialization of the score matrix. In this context, some of the outlier values predicted by the JEDI model may be a factor resulting from something akin to a local minima in the data factorization. Another direction for future work would be to explore how individual subject estimations change given particular decompositions, and how to determine which is closest to the 'true' developmental value.

6.5.3 Application and Relevance

The ability to estimate a child's developmental status from EEG could be beneficial for many applications. In terms of BCI or other rehabilitation strategies in which EEG is featured predominantly, the JEDI model could open up a means to rapidly assess how the intervention was progressing directly from the EEG. This could help alleviate some of the resource intensive requirements for repeated clinical assessments. The present gold standard is to do labor intensive detailed psychometric tests, which are not always available or possible and may suffer from learning effects over repeated testing. The proposed JEDI model thus opens up the potential for a computational biomarker for developmental status clinically. Development of such a biomarker is a sought after, yet elusive, item in trials aimed at assessing clinical interventions to developmental delays.

For EEG based rehabilitation paradigms, it offers an additional use for the recorded EEG data. The data could be used to both infer development and to drive specific tasks of interest, e.g.

MI and MRCP based BCI described in Section 3.3. Furthermore, the JEDI model provides an avenue to more directly link changes found at an electrophysiological level to well-established measurements of development at the psychological level. The proposed JEDI model thus offers multiple benefits in its data-driven approach to inferring a child's developmental status using the common task of a button-press.

6.5.4 The JEDI model Captures EEG Profiles of Interest

Looking at the profiles in Figure 6.5, we can confirm that critical aspects of the button-press task have been captured by the tensor aspect of the JEDI joint decomposition. For example, as this was data averaged over all right-button press trials, we see the first 10 channels (left-brain) in the $[Spatial]$ domain strongly responding for most of the identified components, as compared to the last 10 channels (right-brain). This is expected, as we are examining channels over the central motor cortex, which has contralateral activation with respect to movement (e.g. moving the right arm activates the left motor cortex and vice versa). Additionally, examining the $[Temporal]$ and $[Spectral]$ domains together reveals Factor 2 as potentially capturing the pre-motor desynchronization in the motor cortex of the children, which occurs just before a movement or action (e.g. pressing the button) [165]. The $[Temporal]$ profile indicates the factor is most dominant a full half second prior to the button press, while is almost non-existent during the actual movement. This corresponds to several peaks across the $[Spectral]$ domains frequency spectrum, in particular near 6 Hz and 10 Hz. Both of these peaks correspond to the relative range of alpha/mu, a brain rhythm strongly associated with movement, for children at different developmental levels [171]. These results support that the tensor factorization aspect of the JEDI model was still able to elucidate important information from the EEG in its factor profiles. The JEDI model therefore retains its suitability in characterizing the EEG using the tensor format established in Chapters 4 and 5, while providing an additional means to incorporate developmental information of interest, such as standard clinical measures of development.

6.5.5 Limitations

There were several limitations associated with this work. First, the sample size of this study was relatively small. Due to the exclusive design of the study, over half of the available

children were not included in analysis since they were missing data from either a CCP block, a developmental score or both. Future work should repeat these findings in a larger group to verify the results across a larger population. Additionally, the sample investigated was for typical developing children, and an older childhood population. As development and EEG are heavily linked to age of the child, it is unclear if the JEDI model would be as predictive in children with neurological deficits (e.g. epilepsy or paediatric stroke), or in children who are very young (< 5 years old). Additional investigations should be conducted to demonstrate the model in a variety of paediatric populations.

Another limitation was that the grid search was done prior to the actual analysis, in order to identify a common hyperparameter set. Although computationally expensive, an alternative could be to implement the wide grid search for optimizing hyperparameters at each fold of the analysis. The current work was intended to demonstrate broadly the potential effectiveness of the JEDI model, and as such future investigations could implement this increased complexity. Additionally, the direct projection step has a limitation when estimating new scores from subjects. If a given decomposition includes a near-zero component factor in one of the component matrices, this step will likely result in an error which is outside of the score range. In order to alleviate this concern, we re-ran any individual joint-factorization where the total mean loading value of a component was close to zero.

A trade-off was made in using the characteristic filters as well. The filters require a trivial computational cost (matrix multiplication), but are strictly linear in nature. The 'centering effect' seen in the results may be due to an underlying non-linear relationship between the score data and the EEG. This could be more thoroughly explored in follow-up studies, and potentially addressed using more flexible joint factorization models which could better account for such non-linear structures between the fused data.

Finally, the reported results focus only on the model's ability to predict developmental scores. Further exploration on how to utilize the model to also predict a child's decision between left and right buttons could help bolster the relevance of the data. To approach this, the multi-way model could be evaluated separately for the left and right button data. Then distinct profiles from the left button and right button could be identified, and potentially be used in a binary classification scheme, similar to Chapter 5. This approach would require careful consideration,

however, of how to weight the $[Spatial]$ and $[Spectral]$ modes to help distinguish the relevant button-press information without masking the identified developmental factors. Exploring such a development is an important avenue for future work.

6.6 Conclusions

This chapter outlines using joint-factorization of multi-relational data reflecting child development in a model termed the Joint EEG Development Inference (JEDI) model. The structure of the JEDI model comprised of the fusion of a common EEG task (button press) and a matrix of clinically relevant developmental diagnostic tests (WASI-II, WIAT-IIA, KKQ). Through using multi-way tensors to maintain the inherent structural relationships present in the EEG data, shared developmental information could be gleaned from the SDF. Considerations were given to the relative weights for the tensor-matrix fusion, and regularization factors (L2, L1) through a grid search approach. Results indicated that the JEDI model is able to infer the developmental scores of children, with estimated values falling approximately within the same or adjacent categorical group as the real values. Analysis showed the JEDI model was stable, as similar results were found when using 10-fold cross-validated data and completely unseen EEG data. Furthermore, the JEDI model demonstrated that the expanded tensor framework could effectively characterize important feature profiles in the paediatric EEG data, further supporting the findings of Chapters 4 and 5. The JEDI model therefore is a promising tool with potential application in estimating developmental status of children rapidly using a common EEG task, while still retaining the robust properties of multi-way analysis in characterizing paediatric EEG data.

Chapter 7

Network Analysis of EEG as a Complementary Approach in Characterizing Cognitive Impairment in Preschool Children

This chapter is drawn from the accepted journal manuscript published in *Epilepsy & Behavior* [6].

The key contributions of this chapter are:

- Revealing correlations between network structures and impaired development in children with early onset epilepsy.
- Demonstrating how to utilize network analysis results to classify development from children using EEG.
- Putting forward a potential clinical tool for assessing child development using routine EEGs in very young children.

7.1 Introduction

The JEDI model developed in Chapter 6 presents a promising platform for incorporating alternative measures of development into a multi-way analytic pipeline which characterizes paediatric EEG data. Here, we aim to supplement the JEDI model by developing upon the complementary methodology introduced in Section 2.7, that utilizes functional network analysis to assess cognitive impairment (CI) in children. This chapter examines the alternative analysis in a particular paediatric population prone to CI: children with early-onset epilepsy

(CWEOE), as seen in the MMEC dataset of Chapter 4.

Epilepsy is a complex disease that can have devastating effects on quality of life [173]. CI frequently and severely affects the quality of life of children and their families, and coexists in more than half of children with epilepsy [174–177]. Timely identification of CI, particularly in CWEOE (epilepsy onset at < 5 years of age) is critical because early-life interventions are likely to be more effective [178]. Additionally, it is the period in which childhood epilepsy is most common, and in which the most severe forms of epilepsy occur [179–181]. An estimated 40% of CWEOE have CI [177]. The urgent need for emphasis on early recognition, novel interventions and improved public health strategies for primary and secondary prevention for CI in epilepsy is highlighted in calls to action by prestigious institutions including the International League Against Epilepsy, The Institute of Medicine, and the World Health Organization [182, 183]. Therefore, there is a need to understand the causes of CI and find reliable, affordable and non-invasive markers beyond current standard approaches.

Identification of CI is especially difficult in CWEOE because the gold standard of diagnosis by psychological assessments may not be readily available [184], it is resource intensive, and can be clinically challenging (e.g., introducing potential bias from repeated testing) [184]. Thus, reliable, affordable and rapid CI screening techniques in clinical care are sought after. Such techniques would help focus further medical investigations and resources onto a smaller subgroup, producing efficiency gains and cost savings. Graph network analysis of standard routine clinical EEG recordings is one such potential technique [185, 186].

Analysis of functional EEG networks offers a data-driven methodology for understanding diverse brain conditions through the lens of network (connectivity) properties [55, 187]. As discussed in Section 2.7, functional networks examined as graphs are well-established, and provide advantages in understanding changes in connectivity across the brain [187]. Functional network analysis of the brain can exploit properties like small-world topology, connected hubs and modularity to help inform on the brain's functioning and processing [187]. Insights into epilepsy, including the severity of cognitive disturbances, outcomes of epilepsy surgery, and disease duration have been found to correlate with the extent of changes in these functional networks [56]. Recent work has also found network abnormalities can appear in both ictal and interictal states [56], i.e. during and in between epileptic seizures. This supports that networks

can be distinguished in resting-state EEG [56]. Therefore, functional graph analysis is well positioned as a potential tool to reveal insights into CI in CWEOE.

The aim of this chapter is to explore if reliable EEG network markers which could help effectively screen for CI in CWEOE can be identified. Our hypothesis is two-fold. First, informative network abnormalities can be revealed in CWEOE using graph network analysis on routine clinical rEEGs. Second, identified abnormalities can be integrated into a simple machine learning paradigm to demonstrate predictive capabilities with respect to CI, producing a complementary approach to the goals of the JEDI model from Chapter 6.

7.2 Materials

7.2.1 Dataset

The data used in this study comes from the same MMEC dataset described in Section 4.2.1. Additional information relevant to the current network analysis and epileptic population is provided here for convenience. The complete details on study recruitment and assessments are reported in [178].

Newly diagnosed CWEOE of mixed epilepsy types and aetiologies were recruited as part of a prospective population-based study of neurodevelopment in CWEOE [140] in the Fife and Lothian areas near Edinburgh, U.K. Parents gave approval for use of the standard, resting-state, awake 10-20 EEG their child had as part of their routine clinical care. If a child had multiple EEGs, only the first EEG was used in analysis to avoid biasing results toward children with multiple recordings. Additionally, it allowed similar selection of resting-state recordings across all children, e.g. awake resting-state. As such, no EEG recordings of sleep were analysed in this work.

Table 7.1 provides the demographic and clinical features for the CWEOE which were included in this study. Given the broad anti-epileptic drug (AED) therapies and aetiologies present in Table 7.1, potential interactions from AED load or specific aetiology were examined with respect to the designated CI classes (e.g. normal, mild/moderate, severe CI). Using a non-parametric version of the two-way ANOVA (Friedman's test [188]) on data from Table

	Normal ($n = 31$)	Mild/Moderate CI ($n = 7$)	Severe CI ($n = 13$)
Age in months (SD)	36.18 (19.87)†	26.76 (17.06)	20.37 (18.56)†
Male:Female Ratio	20:11	6:1	6:7
Ethnicity			
Asian	2 (6%)	–	1 (8%)
Black	–	1 (14%)	–
White (U.K./European)	29 (94%)	6 (86%)	12 (92%)
Antiepileptic Drugs			
None	3 (10%)	1 (14%)	–
Monotherapy	26 (84%)	6 (86%)	9 (69%)
Polytherapy	2 (06%)	–	4 (31%)
Focal Seizures	12 (39%)	3 (43%)	4 (31%)
Generalized Seizures	18 (58%)	2 (28.%)	9 (69%)
Generalized and Focal	1 (3%)	2 (28.5%)	–
Epilepsy aetiology			
Cryptogenic	3 (10%)	1 (14%)	5 (38%)
Idiopathic	24 (77%)	4 (57%)	1 (8%)
Symptomatic	3 (10%)	2 (29%)	7 (54%)
Unknown	1 (3%)	–	–
Cognitive z-score (SD)	-0.05 (0.66)	-1.41 (0.20)	-2.9 (0.27)

Table 7.1: Demographic and clinical feature information of patients, grouped by CI classes of normal, mild/moderate CI, and severe CI. Significant differences between groups with respect to age are indicated by a † (Kruskal-Wallis with post-hoc Mann-Whitney U; $H = 6.4697$, $p < 0.05$, with mean ranks of 30, 23.7143, and 17.6923 for Normal, Mild/Moderate CI and Severe CI respectively.)

7.1, we revealed no significant interactions between any AED load or specific aetiology with respect to any CI classes. This in turn suggests that the results identified via the proposed network analysis below are likely driven mainly by cognitive phenomena, as opposed to epileptic syndrome or AED load effects.

A retrospective analysis was done on the 32-channel, unipolar routine EEG [10], described in Section 4.2.1 as the MMEC dataset. Analysis was blinded to any treatment or seizure frequency information. Participants underwent cognitive assessment with age-appropriate standardized tools, e.g. Bayley Scales of Infant and Toddler Development- Third Edition (Bayley-III) and Wechsler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI-III). Children who scored within $\pm$ 1 standard deviation (SD) of the normative mean were defined as normal, -1 to -2 SD as having mild/moderate CI, and < -2 SD as having severe CI. The cognition scores from Bayley-III and WPPSI-III tests were converted into a normalized standard score measure, included at the bottom of Table 7.1. Clinical details were collected by members of the research team using a standardized proforma by direct interview of care-givers, medical records

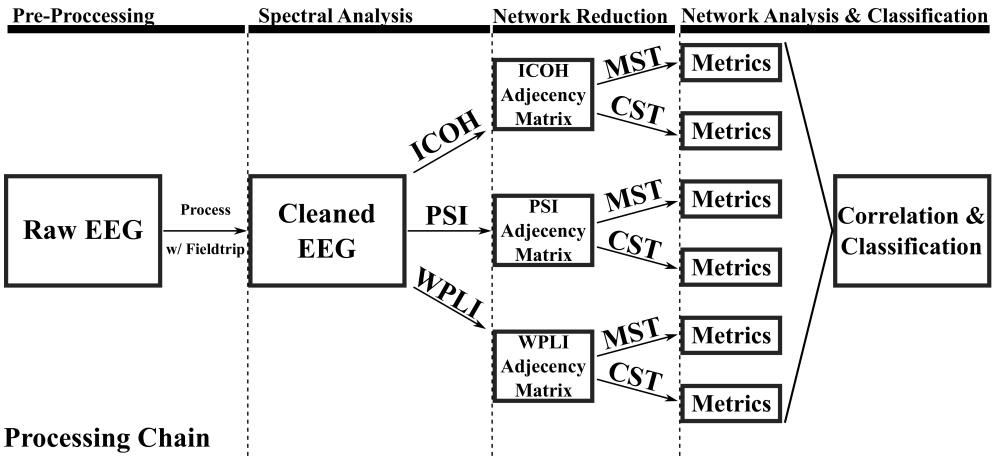

Figure 7.1: Flowchart of the data processing chain for an individual child. ICOH = Imaginary part of coherency, PSI = Phase-slope index, WPLI = Weighted phase-lag index, MST = Minimum Spanning Tree, CST = Cluster-Span Threshold. Metrics refer to set of graph measures outlined in Figure 7.5.

and, where possible, patients themselves when they attended for clinical and/or research study assessment.[1]

7.3 Methods

The data processing pipeline for each child is summarized in Figure 7.1.

7.3.1 Pre-processing

EEG recordings were pre-processed in Matlab using the Fieldtrip toolbox outlined in Section 2.8.1 [58]. The EEG had a sampling rate of 511 Hz from the clinical recording equipment. Recordings were re-referenced to a common average reference, and bandpass filtered between 0.5-45 Hz in Fieldtrip. The resting-state data was split into non-overlapping, two second long sub-trials; long enough to pick up resting-state network activity, while still fitting at least one full period of the lowest included frequency.

Prior to data processing, seizure activity in the EEGs were confirmed by clinicians. Whole trials which contained seizure activity were excluded from the analysis, rather than excluding

[1]The research team was based out of the Royal Hospital for Sick Children and Muir Maxwell Epilepsy Centre in Edinburgh. Researchers and clinicians in the team included Matthew Hunter, Michael Yoong, Krishnaraya Kamath Tallur, Jay Shetty, Ailsa McLellan and Richard F.M. Chin.

only sections of trials with evident seizure activity. This helped guarantee that all network trials were derived from a minimum of two continuous seconds of seizure-free EEG.

Standard EEG artefacts were rejected using the multi-step approach with manual and automatic rejection described in Section 4.3.1. All trials containing EEG artefacts were excluded from analysis. For subjects, we averaged across all trials at each frequency band, to help reduce potential bias and variance resulting from our selection of a shorter analysis window.

A 2-Hz narrow band approach was used in analysis of clean EEG data, similar to work done by Miskovic *et al.* [12]. Segmenting the frequency range into these narrow bands (e.g. 1-3 Hz, 3-5 Hz, etc.) provided a data-driven approach to interrogate networks across subjects. The *a priori* nature of the investigation avoided attempts at equivocating the (likely heterogeneous) impact of epilepsy, development, medication, etc. on each child's spectral EEG composition. While such narrow bands may eschew some physiological interpretations by not adhering to classical frequency bands, the narrow bands promoted identification of mainly robust, common network abnormalities across the heterogeneous CWEOE population. This helps also to account for the migration and dynamic changes in the frequency bands seen throughout early childhood [12, 22]. If significant network abnormalities were identified in these narrow frequency bands (after correction for multiple comparisons, age and spurious correlations) then the identified results were likely representative of a strong effect.

7.3.2 Network Coupling Analysis

This study selected several measures of dependencies in EEG recordings, created graph networks based on these measures and characterized the created networks to identify candidate biomarkers for classification and identification of CI in CWEOE. A summary of the network coupling analysis used to measure dependencies in the EEG are described here for clarity.

Functional EEG network graphs outline the 'functional links' between any pair of EEG channels *i* and *j*, providing a description of brain connectivity. Links are established through measures of interdependency between signals s_i and s_j [185]. A common measurement for examining this interdependency is the cross-spectrum function $S_{ij}(f)$ [185, 189, 190]. Formally, let $x_i(f)$ and $x_j(f)$ be the complex Fourier transforms of the time series signals

s_i and s_j for any pair (i, j) of EEG channels. Then the cross-spectrum can be calculated as

$$S_{ij}(f) \equiv \langle x_i(f) x_j^{\star}(f) \rangle \tag{7.1}$$

where * indicates the complex conjugation, and $\langle\rangle$ refers to the expectation value (also written as $E\{\}$) [190].

This study investigates three connectivity analysis methods building from the cross-spectrum, namely: (1) the imaginary part of coherency (ICOH) [190], (2) phase-slope index (PSI) [191], and (3) weighted phase-lag index [192, 193].

Imaginary Part of Coherence (ICOH)

ICOH is a standard measure in functional network analysis [190]. Coherency is defined as the normalized cross-spectrum [190]:

$$C_{ij}(f) \equiv \frac{S_{ij}(f)}{(S_{ii}(f) S_{jj}(f))^{1/2}} \tag{7.2}$$

The imaginary part of coherency is therefore defined as [190]

$$ICoh_{ij}(f) \equiv Im\{C_{ij}(f)\} \tag{7.3}$$

where $Im\{\}$ refers to taking the imaginary part of a value, in this case of the complex coherency measure. ICOH is well documented, and has been shown to provide direct measures of true brain interactions from EEG while eliminating self-interaction and volume conduction effects [190]. A weakness of ICOH, however, is its dependence on phase-delays, resulting in identifying functional links only at specific phase differences between signals s_i and s_j, while completely failing for others [192–194].

Phase Slope Index (PSI)

The PSI [191] was selected as a complementary alternative to ICOH for analysis. Formally, the PSI is defined as

$$\Psi_{ij}(f) = Im\{\sum_{f \in F} C_{ij}^{\star}(f) C_{ij}(f + \delta f)\} \tag{7.4}$$

where $C_{ij}(f)$ is as defined in equation (7.2), δf is the frequency resolution, and $f \in F$ is the set of frequencies over which the phase-slope is calculated. (See [191] for details.)

In practice, the PSI examines causal relations (temporal order) between two sources for signals of interest, e.g. s_i and s_j [191]. PSI exploits the phase differences between the sources to identify the 'driving' versus 'receiving' relationship between the sources [191]. Their average phase-slope differences are used to identify these functional links [191]. Importantly, unlike ICOH, the PSI is equally sensitive to all phase differences from cross-spectral data [191]. However, the PSI equally weights contributions from all phase differences, meaning even small phasic perturbations in EEG are equal to the (defining) large perturbations.

Weighted Phase Lag Index (WPLI)

The weighted phase-lag index (WPLI) was included as a third comparative measurement for analysis [192, 193]. The standard phase-lag index (PLI) [192] is a robust measure derived from the asymmetry of instantaneous phase differences between two signals, resulting in a measure which is less sensitive to volume conduction effects and independent of signal amplitudes [192]. The PLI is formally defined as [192, 193]:

$$\Theta_{ij} \equiv |\langle sign(Im\{S_{ij}(f)\})\rangle| \tag{7.5}$$

where $sign(\cdot)$ is the positive or negative sign, and $Im\{S_{ij}(f)\}$ is the imaginary part of the cross-spectrum[1]. The PLI ranges between 0 and 1, where a PLI of zero indicates no coupling (or coupling with a specific phase difference; see [192] for details), and a PLI of 1 indicates perfect phase locking [192]. The PLI's sensitivity to noise, however, is hindered as small perturbations can turn phase lags into leads and vice versa [193].

A weighted version of the PLI was introduced (weighted PLI; WPLI) [193] to counter this effect. The WPLI adds proportional weighting based on the imaginary component of the cross-spectrum [193]. The proportional weighting alleviates the noise sensitivity in PLI, and is formally defined as [193]:

$$\Phi_{ij}(f) \equiv \frac{|\langle |\mathcal{S}| sign(\mathcal{S})\rangle|}{\langle |\mathcal{S}|\rangle} \tag{7.6}$$

[1]Note that ICOH in equation (7.3) reflects the imaginary part of the *normalized* cross-spectrum, while the standard cross-spectrum is used here

where $\mathcal{S} = Im\{S_{ij}(f)\}$. The WPLI, like the PSI, helps capture potential phase-sensitive connections present in the rEEG networks from another perspective.

7.3.3 Adjacency Matrices and Sub-Networks

The functional connections found for the ICOH, PSI, and WPLI measures represented as adjacency matrices are provided in the analysis below. A set of adjacency matrices for a representative normal and impaired cognition child in the range of 5-9 Hz are included in Figure 7.2 and Figure 7.3, respectively. The 5-9 Hz frequency range was chosen as the example range to cover the span where the classic alpha frequency band would likely reside for the very young children in this dataset [12, 22].

Constructing and comparing graphs of functional EEG networks built using the adjacency matrix can lead to certain biases in the network analysis [195–197]. To avoid this issue, we used two methods for defining unbiased sub-networks to represent the functional EEG for comparison and analysis: the Minimum Spanning Tree (MST) [195] and the Cluster-Span Threshold (CST) [198].

The MST is an acyclic, sub-network graph which connects all nodes (electrodes) of a graph while minimizing link weights (connectivity strength) based on applying Kruskal's algorithm on the weighted network [195, 199]. In brief, the algorithm first orders the link weights in a descending manner, i.e. from strongest to weakest connectivity [195]. The MST is then constructed by starting with the largest link weight and adding the next largest link weight until all nodes, *N*, are connected in an acyclic sub-network with a fixed density of $M = N - 1$ [195]. After construction of the sub-network, all weights are assigned a value of one [195]. In this manner, the MST is able to efficiently capture a majority of essential properties underlying a complex network in an unbiased sub-network [195].

Exploiting the properties of the MST is a relatively recent technique, presented in contemporary publications exploring brain networks [195]. However, the MST naturally leads to sparse networks in the data due to its acyclic nature and, in some occasions, more dense networks may be preferable. Thus, real brain network information is potentially lost in MST based EEG graph analysis [200].

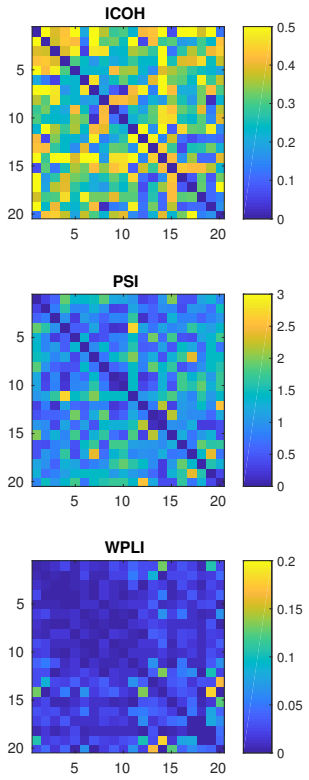

Figure 7.2: Adjacency matrices for a representative 'normal cognition' child calculated by ICOH, PSI and WPLI between the 5-9 Hz frequency range. The top colored graph represents the ICOH's adjacency matrix, where connectivity estimates between all pairs of EEG channels are visualized using the 0-0.5 (arbitrary units) derived from the ICOH network coupling analysis. For example, the estimated connectivity for ICOH between an EEG electrode and itself would be 0, as seen on the diagonal of the ICOH adjacency graph. Similarly, the middle graph depicts the adjacency matrix found from the PSI coupling analysis. Importantly, the range of the colorbar for the PSI adjacency matrix reflects the values found via the PSI, which is calculated separately from either ICOH or WPLI. The bottom graph shows the adjacency matrix for WPLI coupling analysis, where the scale again is distinct from ICOH and PSI graphs. Relative percentage of high estimated connectivity found via the three separate network coupling measures provides insight into potential advantages/disadvantages of these techniques.

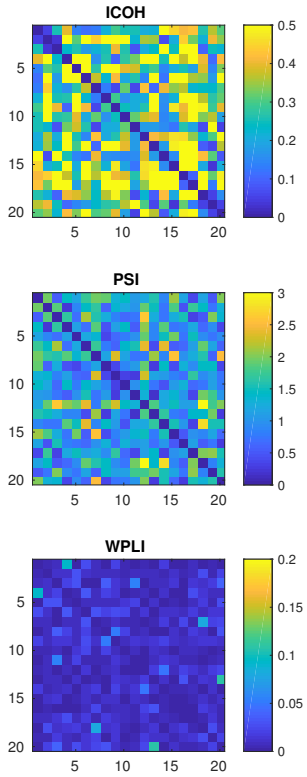

Figure 7.3: Adjacency matrices for a representative 'impaired cognition' child calculated by ICOH, PSI and WPLI between the 5-9 Hz frequency range. Similar to Figure 7.2, each color scale (arbitrary units) shows the estimated connectivity using the particular network coupling analysis technique, e.g. ICOH, PSI or WPLI. A notable reduction is seen in the estimated connectivity across the three adjacency matrices, as compared to those presented in Figure 7.2.

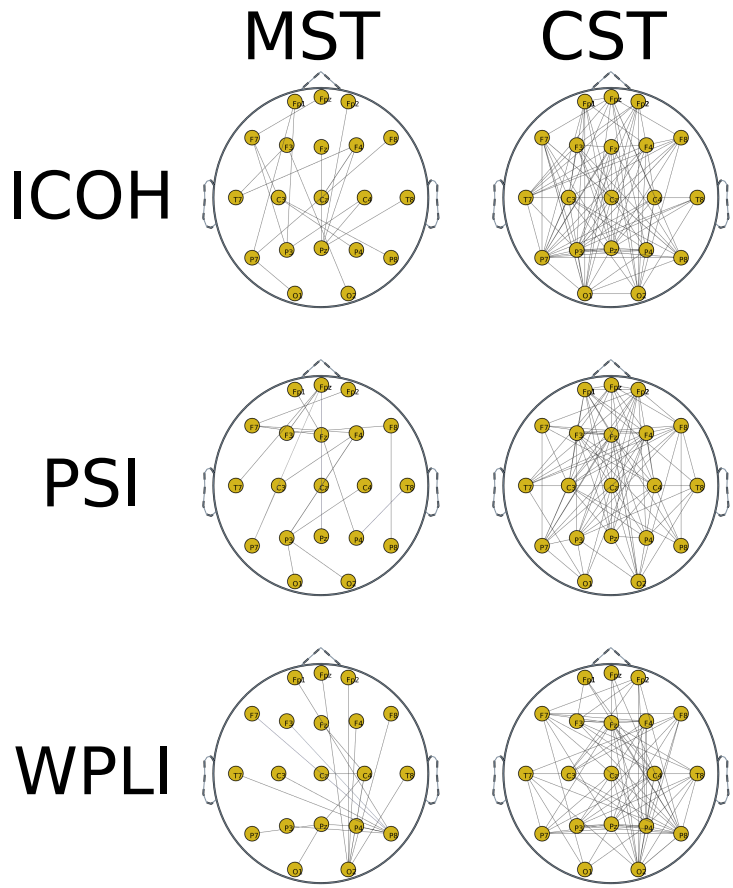

Figure 7.4: Illustrative examples of the MST and CST sub-network graphs of ICOH, PSI and WPLI for a randomly selected child. EEG channels are displayed as nodes, with functional connections displayed for each combination of sub-network and connectivity measure.

By contrast, the CST creates a similar sub-network to the MST, but balances the proportion of cyclic 'clustering' (connected) and acyclic 'spanning' (unconnected) structures within a graph (for details see [198]). This balance thus retains naturally occurring 'loops' which can reflect dense networks without potential information loss [200]. Figure 7.4 illustrates a topographical example of EEG channels connected via MST and CST networks for a randomly selected child. Differences in sparsity between the acyclic MST and the cyclic CST sub-networks can readily be seen in Figure 7.4. Both the MST and CST are binary sub-networks and consequently have advantages over weighted networks which can be affected by spurious connections and link density effects [195, 198, 200].

For each combination of sub-networks and connectivity definitions above (e.g. MST-ICOH, CST-ICOH, MST-PSI, etc.) four network metrics were investigated for correlation to the

cognition standard score measures. To help reduce potential selection bias, network metrics for analysis were agreed upon a priori. Metrics were chosen to account for distinct network properties (e.g. the shape of the network, the critical connection points in the network, etc.) with (relatively) little inter-correlation. Due to the natural exclusion/inclusion of cycles, the network metrics differ between the MST and CST, respectively. However, all metrics across sub-networks were selected to be comparable regarding what network properties were being measured. Pictorial examples of the selected network metrics, alongside short definitions, are given in Figure 7.5.

7.3.4 Statistical Analysis

Statistical analysis was done using Matlab 2015a. Correlation between individual network metrics and the cognition standard score was measured using Kendall's tau (τ) [201]. Kendall's τ calculates the difference between concordant and discordant pairs [201, 202], and is a strong choice for describing correlation in ordinal or ranking properties. In this work, the normalized cognition standard scores' relative rankings serve as the ordered data explored using the τ correlation. The design of Kendall's τ is also relatively robust to false positive correlations from data outliers [201, 202], providing additional mitigation to spurious correlations in the results. Furthermore, as Kendall's τ is a non-parametric hypothesis test it did not rely on any underlying assumptions about the distribution of the data. Therefore our correlation analysis was robust to any potential ceiling, floor or skewed distribution effects present in the reported cognition standard score measures.

Correlation trends are reported both as uncorrected $p < 0.05$ values, and with multiple comparison (Bonferroni) corrections, similar in style to previous literature [203]. For each frequency bin (2-Hz wide) and network, we compared and corrected for the 4 separate graph measures using the Bonferroni technique (i.e. we set $p = 0.05/4 = 0.0125$ as the threshold for significance). Dependency was assumed across the small 2-Hz frequency bins, similar in principle to [203], and as such we do not include the frequency bins in the Bonferroni correction. Correlations which are found to be potentially significant under this assumption are indicated by the † symbol for Bonferroni corrections.

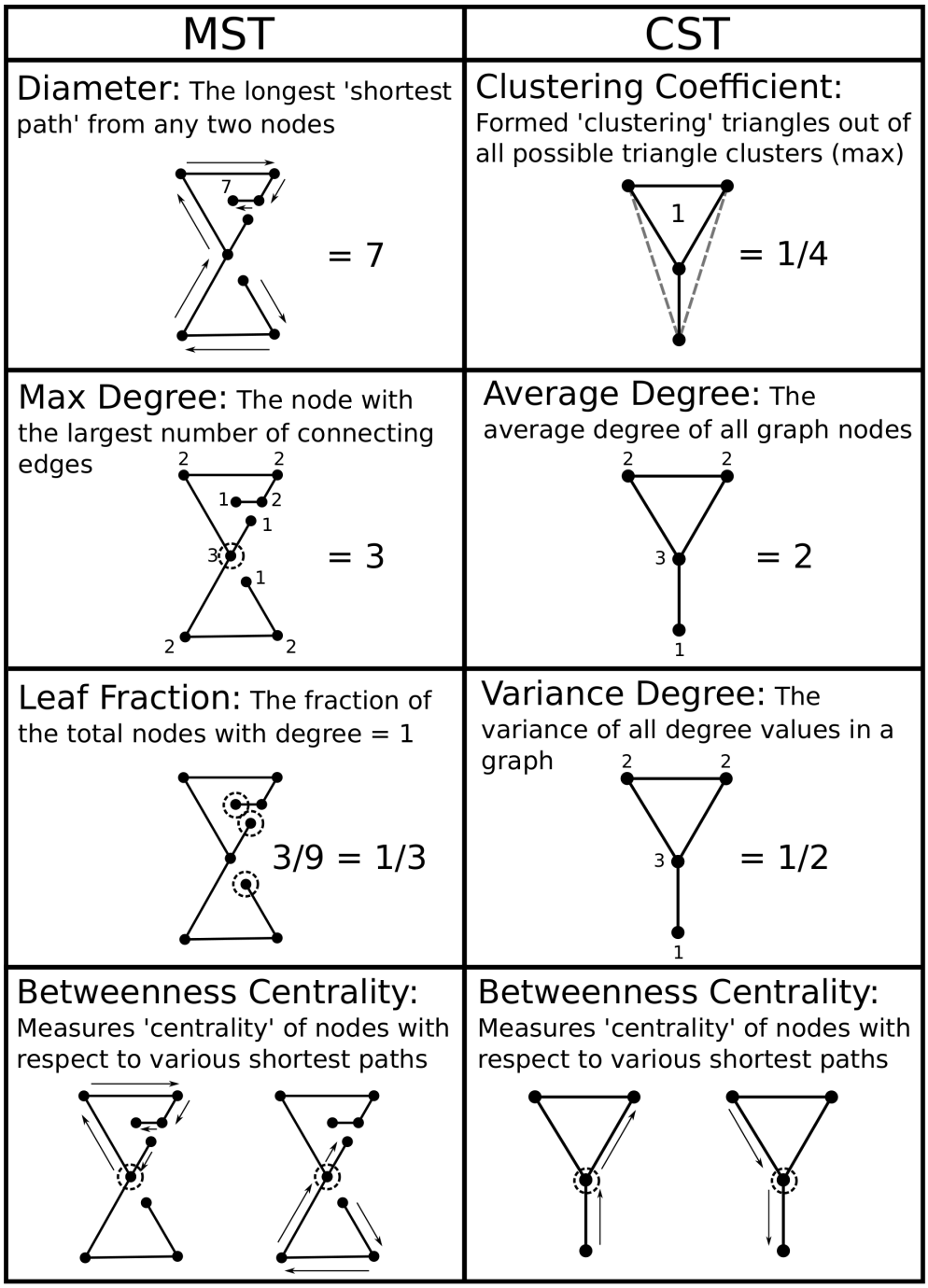

Figure 7.5: Illustration of all graph analysis metrics for the Minimum Spanning Tree (MST) and Cluster-Span Threshold (CST) networks using simple example graphs. Nodes (dots) represent EEG channel electrodes. Edges (lines) represent functional interactions between EEG channels identified by a connectivity measure, e.g. ICOH/PSI/WPLI.

Multi-class Classification Cost Matrix

		CI-Predicted Class		
		Normal	Mild/Mod.	Severe
CI-True Class	Normal	0	2.5	2.5
	Mild/Mod.	5	0	1
	Severe	5	1	0

Table 7.2: Weighted cost matrix for misclassification of cognitive impairment (CI) for normal (± 1 SD), mild/moderate (-1 to -2 SD) and severe (< -2 SD) classes. Rows represent true class labels, with columns as the predicted classification labels.

7.3.5 Classification

A multi-class classification scheme was devised using the WEKA toolbox [61, 144]. Class labels of *normal*, *mild/moderate CI*, and *severe CI* were applied.

Primary feature selection included all correlations identified by the statistical analysis, thereby promoting interpretation in the retained network features. Then, a second feature selection phase using nested 5-fold cross-validation selected prominent features via bi-directional subspace evaluation [204]. Within this nested cross-validation, features identified as important in $> 70\%$ of the folds were selected for use in classification.

Due to the natural skew of the data (towards normalcy), and the context of the classification problem (e.g. misclassifying different classes has various implications), a cost-sensitive classifier was developed [205]. In order to properly develop such a classifier, an appropriate cost matrix needed to be identified. Using guidelines outlined in the literature [205], the cost matrix in Table 7.2 was developed, with predicted classes represented on the rows and real classes given on the columns.

The defined matrix satisfies several key concerns in multi-class cost-matrix development [205]. The weights on misclassification were carefully selected to reflect probable clinical concerns in classification with guidance from paediatric neurologists[1]. The cost for incorrectly classifying an impaired child as normal was twice as heavy compared to misclassifying a normal child into either impaired group. This was still significantly more punishing than if impairment was correctly identified but misclassification occurred in determining between

[1] Dr. Richard F.M. Chin and Dr. Jay Shetty

mild/moderate or severe impairment. These weighted values prioritized accurately including as many 'true positive' CWEOE with CI first, i.e. increasing sensitivity, followed by a secondary prioritization on being able to discern the level of CI. These boundaries provide a more clinically relevant classification context in the analysis.

Using the selected features and developed cost-sensitive matrix, a nested 5-fold cross-validation trained a simple *K*-Nearest Neighbour (KNN) classifier, with $K = 3$ neighbours and Euclidean distance to minimize the above costs. By demonstrating our proof-of-concept results with a simple classifier first, e.g. KNN, we aimed to highlight that network response found from our analysis pipeline was likely robust. A repeated 'bagging' (Boostrap Aggregation [206]) approach was used to reduce variance in the classifier at a rate of 100 iterations/fold. Results were evaluated upon their overall classification accuracy and total penalty costs (e.g. sum of all mistakes based on the cost matrix).

Random selection (RS) and single class (SC) classification were included for comparison, as described at the end of Section 4.3.6. For the RS classificaiton scheme, a chance level for each RS class is used to assign the predicted label at random, based on the distribution of subjects in each class. For the SC scheme, classification accuracy and misclassification penalties are calculated based on the presumed (single) class assignment. This study looked at SC classification for each class label, and have reported comparisons to each possible single class selection in the SC classification.

7.4 Results

Of 64 children enrolled into the parent study, 13 were excluded from the current analysis due to corrupted EEG data and inconsistent or incompatible EEG acquisition parameters. There were data available for analysis on 51 children (32:19 male-to-female ratio, mean age and SD of 30.85 ± 20.08 months). On average approximately 455 ± 325 two second trials were used for each child in the analysis, totalling 15.16 ± 11.87 minutes of resting-state EEG data for each child. Thirty-one children had normal cognition, 7 had mild/moderate CI, and 13 had severe CI.

MST analysis of cognition standard score measures			
Network Type	Network Measurement	Frequency Range(s) (Hz)	Correlation ($\bar{\tau} \pm SD$)
ICOH	Diameter	–	–
ICOH	Maximum Degree	–	–
ICOH	Leaf Fraction	–	–
ICOH	Betweenness Centrality	13-17 Hz	-0.231 ± 0.001
PSI	Diameter	9-19 Hz	$0.239 \pm 0.032^{\dagger *}$
PSI	Maximum Degree	11-13 Hz	$-0.232 \pm 0.000^{*}$
PSI	Maximum Degree	15-17 Hz	$-0.258 \pm 0.000^{\dagger *}$
PSI	Maximum Degree	21-23 Hz	-0.219 ± 0.000
PSI	Leaf Fraction	11-13 Hz	-0.201 ± 0.000
PSI	Leaf Fraction	15-19 Hz	-0.246 ± 0.003
PSI	Betweenness Centrality	9-13 Hz	$-0.218 \pm 0.012^{*}$
PSI	Betweenness Centrality	17-19 Hz	$-0.259 \pm 0.000^{\dagger *}$
WPLI	Diameter	–	–
WPLI	Maximum Degree	29-31 Hz	$-0.310 \pm 0.000^{\dagger *}$
WPLI	Leaf Fraction	–	–
WPLI	Betweenness Centrality	23-25 Hz	0.223 ± 0.000

Table 7.3: Summary of Kendall's τ correlation trends between various graph metrics and the cognition standard scores using the Minimum Spanning Tree (MST). For all values $|\tau|$ was between 0.201 and 0.310; mean $= 0.239 \pm 0.0278$ and uncorrected $p < 0.05$. Significant values across contiguous narrow-band frequencies have been grouped together for ease of interpretation.

† *Significant with Bonferroni correction at the level of frequencies.*

* *Significant after partial correlation correction to age of subjects, via modified τ with uncorrected $p < 0.05$.*

7.4.1 Correlation Analysis of the Network Metrics

Each combination of functional link analysis (ICOH/PSI/WPLI) and sub-network selection (MST/CST) techniques uncovered correlations between at least one network metric outlined in Figure 7.5 and the cognition standard score measures. A summary of the significant correlations between the MST metrics and the standard scores are shown in Table 7.3. All MST correlations were in the medium to high frequency range, $9 - 31$ Hz, with no significant results in lower frequencies. Activity above approximately 9 Hz is outside of the expected range for the delta, theta and alpha bands in young children [22,207]. Sets of contiguous frequency bands with significant correlations were found in the ICOH and PSI connectivity measures, and are reported together as a single frequency range. Overlapping correlations retained at significant levels after partial correlation correcting for age are also reported for the MST using a modified Kendall's τ.

CST analysis of cognition standard score measures			
Network Type	Network Measurement	Frequency Range(s) (Hz)	Correlation ($\bar{\tau} \pm SD$)
ICOH	Clustering Coefficient	15-17 Hz	$-0.290 \pm 0.000^{\dagger *}$
ICOH	Average Degree	–	–
ICOH	Variance of Degree	13-15 Hz	-0.200 ± 0.000
ICOH	Variance of Degree	21-23 Hz	-0.203 ± 0.000
ICOH	Betweenness Centrality	11-13 Hz	$-0.273 \pm 0.000^{\dagger *}$
ICOH	Betweenness Centrality	15-17 Hz	-0.241 ± 0.000
ICOH	Betweenness Centrality	19-21 Hz	-0.203 ± 0.000
PSI	Clustering Coefficient	–	–
PSI	Average Degree	13-15 Hz	-0.210 ± 0.000
PSI	Variance of Degree	15-17 Hz	$-0.277 \pm 0.000^{\dagger *}$
PSI	Variance of Degree	21-23 Hz	-0.217 ± 0.000
PSI	Betweenness Centrality	5-7 Hz	$0.204 \pm 0.000^{*}$
PSI	Betweenness Centrality	15-17 Hz	-0.248 ± 0.000
WPLI	Clustering Coefficient	1-3 Hz	$-0.236 \pm 0.000^{*}$
WPLI	Clustering Coefficient	17-19 Hz	$0.287 \pm 0.000^{\dagger *}$
WPLI	Average Degree	–	–
WPLI	Variance of Degree	1-3 Hz	$-0.236 \pm 0.000^{*}$
WPLI	Betweenness Centrality	–	–

Table 7.4: Summary of Kendall's τ correlation trends between various graph metrics and the cognition standard scores using the Cluster-Span Threshold (CST). For all values $|\tau|$ was between 0.201 and 0.290; mean $= 0.237 \pm 0.033$, and uncorrected $p < 0.05$. Significant values across contiguous narrow-band frequencies have been grouped together for ease of interpretation.

† *Significant with Bonferroni correction at the level of frequencies.*

* *Significant after partial correlation correction to age of subjects, via modified τ with uncorrected $p < 0.05$.*

Similarly, significant correlations between the CST metrics and the cognition standard scores are shown in Table 7.4. Several significant CST metrics exist in the lower frequency range (< 9 Hz), indicating a potential sensitivity of the CST to lower frequencies. No sets of continuous frequency bands were discovered, but several sets were trending towards this phenomenon within ICOH. Several correlations in the CST remained significant after partial correlation correction for age from the modified τ at lower frequencies, indicating a potential additional sensitivity in the CST sub-network.

Both the MST and CST demonstrate similar or higher sensitivity (i.e. significant correlations) in the phase-dependent measures (PSI, WPLI) compared to the ICOH measure.

Confusion Matrix from Classification Results

		CI-Predicted Class		
		Normal	Mild/Mod.	Severe
CI-True Class	Normal	**26**	*2*	*3*
	Mild/Mod.	2	***3***	*2*
	Severe	1	*5*	***7***

Table 7.5: Resulting confusion matrix from the 5-fold cross-validated, cost-sensitive classification scheme for all $n = 51$ children based on costs in Table 7.2. Rows represent true class labels, with columns as the predicted labels from the classification. Bold values along the diagonal show true positive classification results, where actual and predicted cognitive classes were accurately identified. Italicized values indicate children predicted to have CI, i.e. mild/moderate or severe class, by the classification scheme.

7.4.2 KNN Classification of the CST

Based upon the CST's potential sensitivity, a preliminary classification scheme assessed the possible predictive qualities of the CST network metrics in identifying CI classes. The relative quality of the classifications are examined using classification accuracy and total 'cost' (i.e. penalty for misidentification) [205].

The subset of CST metrics for classification were identified from the significant correlations above and chosen via the cross-validated feature selection scheme given in Section 7.3.5. The chosen features included five network metrics across the three connectivity measures of ICOH, PSI and WPLI.

From ICOH, the identified feature subset selected included the betweenness centrality at ranges 11-13 and 19-21 Hz, as well as the clustering coefficient in the 15-17 Hz range. From the PSI, the feature subset included the average degree in the 13-15 Hz range. From the WPLI, the variance degree from 1-3 Hz was selected. These results indicate specifically which network metrics, from a machine-learning perspective, contributed the most information for building an accurate classification model. The classifier was trained specifically using these 5 key metrics. An illustrative example of these 5 selected network metrics (e.g. features) are shown in Figure 7.6 as scatter plots.

The resulting confusion matrix from the 5-fold cross-validated, cost-sensitive classification analysis is seen in Table 7.5.

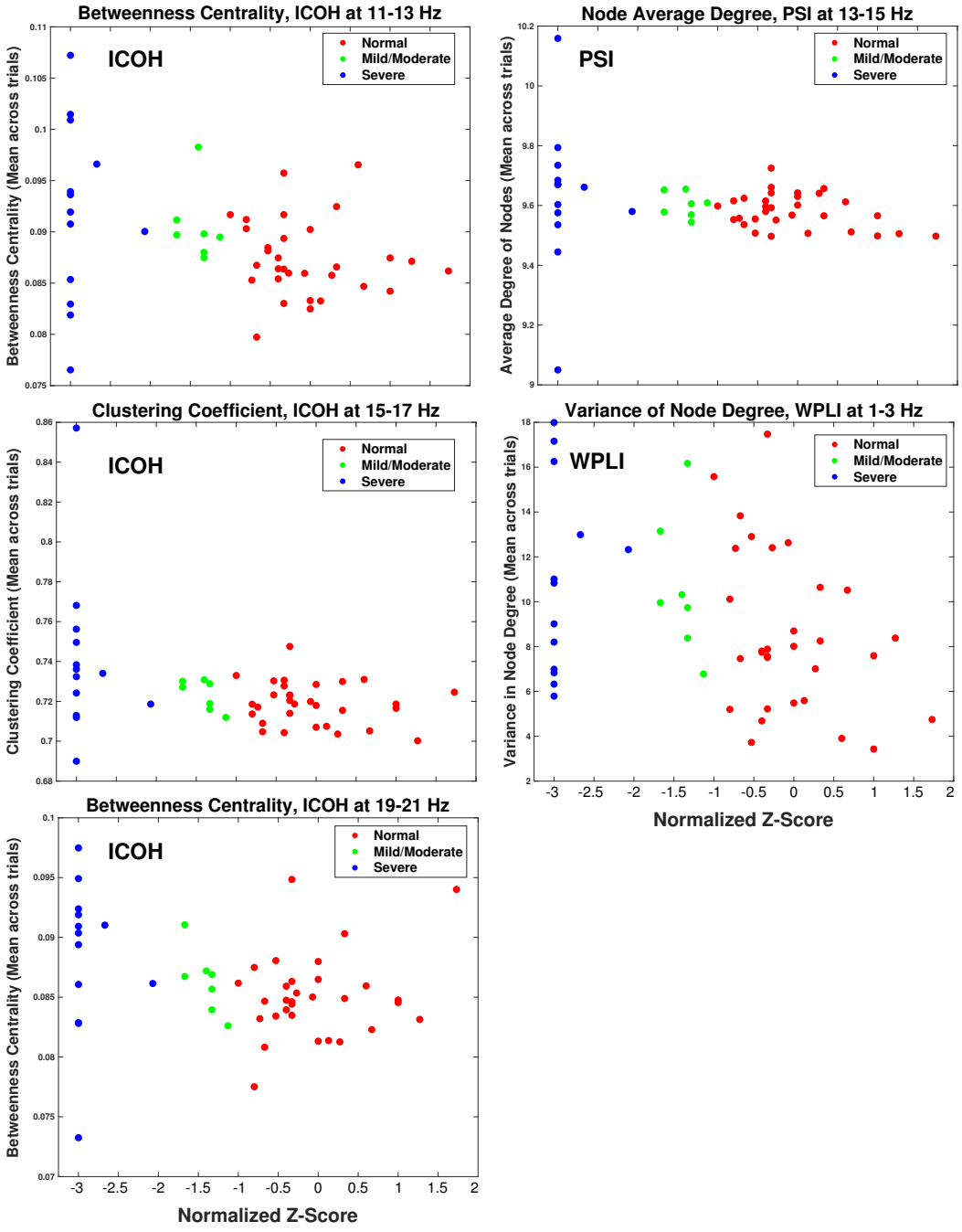

Figure 7.6: Scatter plot displaying the distribution of children for each of the 5 features used in training the KNN classification. Each panel displays network values on the y-axis, with the normalized cognition standard score (z-score) on the x-axis. Children classified into normal, mild/moderate CI and severe CI classes are displayed in red, green and blue respectively.

The overall classification accuracy was defined as the number of true label classes correctly predicted by the classifier, e.g. the true positive diagonal of Table 7.5. Presently, approximately 36 of the 51 children's cognitive class (e.g. normal, mild/moderate CI, severe CI) were correctly predicted, giving a total accuracy for the classifier at 70.6%. Using Table 7.2, an overall 'cost-penalty' value was calculated at 38 points, based on the children who were misclassified, i.e. their cognitive class was not correctly predicted.

The expected RS classification accuracy is based on the distribution of individuals belonging to each class, i.e. 31, 7 and 13 children for the *normal*, *mild/moderate* and *severe* classes respectively. Accuracy for the RS classification scheme would be expected at 45.4%, with cost-penalties varying depending on misclassification distributions. Using the average misclassification penalty and the percentage of misidentified children (approximately 28 of the 51 subjects), the cost-penalty would be at least 65 points for the RS scheme.

The SC classification scheme assumed all subjects belonged to only a single cognition class in order to calculate the accuracy and misclassification costs under this scheme. As an example, if all children were considered as belonging to the 'normal' cognition class (i.e. naively classified as normal), then exactly 31 of the 51 children (those whose true class actually is 'normal'-the first row of Table 7.5) would be correctly identified. This would give the SC classification scheme an accuracy of 60.8%. Repeating this, the SC classification scheme for mild/moderate and severe classes has classification accuracies of 13.7% (7/51), and 25.5% (13/51) respectively. Similarly, the total cost-penalty for each SC classification would be 100, 90.5 and 84.5 points respectively, using the same procedure and the penalty costs from Table 7.2.

Overall, the results indicate gains in classification accuracy and a reduced total penalty as compared to both RS and SC classification. This is summarized in Table 7.6.

7.4.3 Correlation to rEEG CP Model

The network analysis may be considered as a complementary approach to analyzing paediatric EEG at several levels. First, the network analysis is complementary at the methodological level of how to characterize the paediatric EEG recording, i.e. complementary to the JEDI model developed in Chapter 6. Second, the network analysis could be complementary at the individual

	Classification Scheme			
	Network Analysis	Random Selection	Single (Naive) Class	SC Value
Total Accuracy	70.6% (36/51)	45.4%(≈23/51)	Normal Cognition	60.8% (31/51)
			Mild/Moderate CI	13.7% (7/51)
			Severe CI	25.5% (13/51)
Total Cost Penalty	38 pts	≈65 pts	Normal Cognition	100 pts
			Mild/Moderate CI	90.5 pts
			Severe CI	84.5 pts

Table 7.6: Summary table of overall classification accuracies and total cost penalty for the proposed network analysis, RS classification, and SC (naive) classification. SC classification is split to show overall classification accuracy and cost penalties if all children were assigned as normal cognition, mild/moderate CI or severe CI classes. Total accuracy includes the approximate number of children with true positive predictions, out of total number of children evaluated.

Tensor Factor	ICOH Btw. Cent. 11-13 Hz	ICOH Clust. Coef. 15-17 Hz	ICOH Btw. Cent. 19-21 Hz	PSI Avg. Degr. 13-15 Hz	WPLI Var. Degr. 1-3 Hz
Fct. 1	-0.051	0.005	0.118	-0.012	0.086
Fct. 2	-0.018	-0.021	0.185	0.094	**0.304**
Fct. 3	0.241	-0.017	0.116	**0.330**	-0.145
Fct. 4	-0.133	-0.151	-0.255	-0.104	-0.064
Fct. 5	-0.114	-0.142	-0.030	0.146	0.200
Fct. 6	-0.104	-0.214	**-0.304**	-0.127	-0.158
Fct. 7	0.123	0.188	0.226	0.095	-0.103
Fct. 8	-0.185	**-0.341**	**-0.372**	-0.106	-0.068

Table 7.7: Resulting correlation values (Pearson's r) between the key factors identified for the rEEG via multi-way analysis (Figure 4.4) and the key network metric properties identified in this chapter. Bold values indicate uncorrected correlations which were significant from zero ($p < 0.05$). Correcting for multiple comparisons using a FDR of $p = 0.05$ found no significant correlations existed between the data.

variable level with respect to information gained from the rEEG. That is, provide supplemental information unrelated to the identified data from the multi-way analysis framework. To investigate this complementary nature at the variable level, the 5 key network feature properties from the CST were correlated with respect to the critical rEEG multi-way factors identified in Figure 4.4 in Chapter 4, using Pearson's correlation. Results from the correlation are given in Table 7.7. While some trends in the correlation emerge, no significant correlations from zero are found between the network information and multi-way factors when corrected for multiple comparisons using a False Discovery Rate (FDR) of $p = 0.05$. These results emphasizes the complementary nature of the two methods, highlighting that the information derived from the two analysis schemes reflect different aspects related to cognitive development in children.

7.5 Discussion

The main finding of this study is the development of network analysis methods to identify potential data-driven markers for CI in CWEOE directly from rEEG. Our results revealed a substantial pool of potential network characteristics which might be helpful in identifying CI in CWEOE via several different network analysis and filtering combinations. The breadth of these combinations emphasizes that network analysis of paediatric rEEG is well suited for identifying CI markers in CWEOE. Furthermore, the study complements our findings in Chapter 6, by providing an alternative methodology useful in characterizing cognitive development of children rapidly via paediatric EEG. The automated and quantitative nature of the processing chain, its ability to appropriately predict CI classes, and its use of routinely acquired EEG data also make the proposed methods an attractive proposition for clinical applications.

Flexibility in sensitivity and robustness of particular networks to features of interest is an advantage of this analysis. For instance, the sensitivity of phase-dependent connectivity measures, e.g. PSI and WPLI, was more prevalent compared to standard ICOH. This is not surprising as phase-oriented measures were developed to improve upon phase ambiguities in traditional ICOH measurements [191, 194]. In addition, the sensitivity of PSI in picking up significant correlations can be attributed in part to its equal treatment of small phase differences in leading and lagging signals [191]. Such small phase differences contribute equally in PSI, while counting for proportionally less in the WPLI by definition [192, 193]. By construction, the WPLI results are substantially more robust to noise and small perturbations in phase, through proportionally reflecting phase differences in network connections with appropriate weights, providing results for only large phase differences. Together these measures reflect trade-off choices between sensitivity and robustness available to the network analysis scheme.

Of interest for paediatric populations is the CST's capability to identify low frequency correlations in phase-dependent coherency measures. Both the PSI and WPLI demonstrate sensitivity to lower frequencies, not present in the ICOH or MST, in general. This is critical considering that in preschool children lower frequencies typically contain the bands of interest present in adult EEGs, e.g. the delta/theta/alpha bands [22, 207], as discussed in Section 2.3.3 and throughout Chapter 3. During development these bands shift to higher frequencies [120],

reflecting a large scale reorganization of the endogenous brain electric fields and suggesting a transition to more functionally integrated and coordinated neuronal activity [12]. While significant correlations found in the low frequencies may potentially be spurious, the (low) chance that all such significant findings in the CST were spurious is less detrimental than the potential impact from disregarding these findings altogether. The ability to detect network disruptions potentially present in these critical bands in CWEOE provides high impact value, and the possibility for adjusting potential therapeutic and treatment strategies for clinicians and researchers.

The identified subset of metrics for classification provide additional information. All of the features in the subset reflected distribution measures of hub-like network structures in the brain, relating to the balance between heterogeneity and centrality within the network. The implicated metrics, other than the variance degree, corresponded to measures identifying local, centralized 'critical' nodes in a network. Their negative correlation to the cognition standard scores imply that children with more locally centralized brain networks, and consequently with less well distributed hub-like structures, are more likely to have corresponding cognitive impairment. This is reasonable, since if there exists a small set of central, critical hubs responsible for communication across the brain, disruption of these critical points (e.g. due to seizure activity and/or diffuse damage in the brain) would have severely negative effects on communication passing through these connections. This is also supported by the negative correlation in the variance degree metric in the WPLI. The variance degree can be interpreted as a measure of a network's heterogeneity [208]. As such, the negative variance degree in the low (1-3 Hz) frequency range may reflect stunted cognitive development, as normal maturation is associated with reduced activation in low frequencies [21, 22, 28, 207, 209], implying a decrease in local connectivity and heterogeneity of the networks. This complements the above conclusions, suggesting a sensitivity in the likely well-centralized networks to significant disruptions by epilepsy. The disrupted networks may then be reflected by the continued heterogeneity and local connectivity of low frequency structures in impaired children.

Being able to predict the likely extent of CI using the identified markers could provide an advantageous tool for clinicians. Specifically, being able to pair specific network features to an effective prediction of CI would allow clinicians to retain interpretation of the chosen network

features while providing a tool to quickly and objectively separate similar cases. To this end, the cost-sensitive, simple KNN classifier explored in this work illustrates an early step towards this aim. Evaluating the network-based classifier results show the analysis was successful at two levels. First, the proposed classifier was able to generally identify cognitively normal children from impaired children, when grouping the mild/moderate CI and severe CI classes. This is seen in the first column of Table 7.5 where only three impaired children are misidentified as 'normal cognition', giving a sensitivity of 85%. In other words, 17 of the 20 actual impaired children were correctly identified as belonging to either the mild/moderate or severe CI classes. This demonstrated that the proposed network analysis classifier was largely successful with respect to predicting children with some form of impaired cognition, based on using the standard score definition. Similarly, only five normal children were misidentified as generally impaired (i.e. classified to either the mild/moderate or severe CI classes; top row of Table 7.5), giving a specificity of approximately 84% (26/31) for appropriately identifying children in the range of normal cognition. In addition, the network coupled classifier was able to separate out cases of mild/moderate impairment from severe impairment decently, with $> 50\%$ of impaired children correctly predicted. Thus, the proposed classifier and associated methods provide considerable sensitivity (85%) and specificity (84%) for clinicians in determining potential CI, while still remaining relatively accurate in separating CI according to severity.

Statistical analysis in this manuscript was utilized as a first-pass means to reduce the potential feature space for classification. Through identifying potentially significant networks of interest, the number of features to test in the classification step was substantially reduced. Through the statistical analysis, we were able to select pertinent features from a relevant and manageable feature space. It bears repeating that Kendall's τ was a non-parametric significance test, which means it did not rely on an underlying assumption of a normal (or any other) distribution in the data. Kendall's τ correlation was therefore robust to the apparent flooring effect seen in the severe CI class, as it utilized concordant and discordant pairs. As such, our conclusions drawn from the statistical analysis were unaffected by this phenomena. Future endeavours could refine such features, based on different choices for the statistical analysis. Using a more rigid/flexible analysis could lead to further culling/relaxation of the feature space and provide an adjustable framework for examining network property changes in CWEOE. Other future work could include alternative narrow-band frequency binning and less strict

automated rejection methods. Significant correlations across sets of consecutive (and nearly consecutive) frequency bands indicate likely targets for other potential follow-up studies. Further development of a more complex classification scheme could help improve the second tier discrimination of the proposed classifier, at the level of discerning between the cognitive impairment types (e.g. mild/moderate CI from severe CI). A thorough investigation into incorporating and comparing additional classifiers is also a potential avenue for expansion of this research.

The NEUROPROFILE cohort was an advantageous for our analysis in that formal neuropsychological testing was coupled with EEG recordings, making it ideal for this investigation. However, there are study limitations. Although this novel study examined routine clinical rEEGs used in the diagnosis of incidence cases of CWEOE, the three classes of normal, mild/moderate and severe impairment were unbalanced; this occurred naturally. The majority of the sample was taken from a population-based cohort, and mitigating potential influences from the imbalanced data was taken into account as much as possible when conducting the research, e.g. through cost-sensitive analysis. Imbalanced data is not uncommon, and the unbalanced distribution of CI in the current study reflects findings from a true population-based cohort [140]. Furthermore, trialling this methodology in older children with epilepsy may be an avenue for future studies. Older subjects would provide further insights as to the relationship between aetiology and CI, as well as provide additional replications of the proposed techniques.

7.5.1 Limitations

Within the studied cohort of CWEOE, the epilepsy type and aetiologies were heterogenous. Thus we are unable to determine if the model and methods used have greater or lesser predictive value in specific subsets of epilepsy. Testing in a larger, more homogeneous sample would provide clarification.

A gender disparity was noted within the normal cognition and mild/moderate CI groups. Although this study reflects a true population, further studies are needed to investigate any effects from this phenomena.

Note that the spectral components in the very low frequency narrow band (e.g.1-3 Hz) may not

be fully reliable due to the small epoch length, i.e. two seconds. Information gained from the very low frequency band needs to be interpreted with care, as spurious connections are more likely to be present. Again, the large number of trial epochs averaged for each child helped mitigate the potential for such spurious connections. However, the effect of the epoch length is a point for future work. Examining how changes in epoch length could affect the resulting analysis would help support presented correlations found in the very low frequency range.

We recognize a limitation in our assumption of dependency between the frequency bins. While there is likely a strong local family dependency between the narrow bins, the endpoints on our frequency spectrum may not have as strong of a relation. Therefore, significance at these level should be considered carefully as they are more likely to be a false positive. However, the robust nature of τ and our choice of features from a machine-learning perspective help to moderate potential impacts from this assumption on our results.

The use of a data-driven, narrow band approach in our analysis had a trade-off of not using patient-specific frequency ranges for each child. Future studies could be done to investigate how individualized frequency bands (i.e. individual alpha frequencies) could be aligned, interpreted and correlated when assessing network abnormalities in the CWEOE population.

7.6 Conclusion

This chapter introduced a novel processing chain based on network analysis for identifying CI in CWEOE. The results demonstrated that the proposed network analysis framework was capable of identifying critical structures in the functional networks of CWEOE who had CI. This work illustrated the potential predictive abilities of the identified network analysis markers using a simple classification scheme. Together these results highlight a separate methodology, complementary and supportive to the multi-way framework developed in Chapter 6, which can characterize cognitive development in children using rEEG data.

Chapter 8

Conclusions and Future Work

8.1 Discussion

This thesis introduced a range of novel multi-way analytic methods directed at characterizing aspects of paediatric EEG in order to improve translation of EEG-driven technologies for children. The work successfully applied the multi-way framework to a variety of paediatric EEG scenarios including:

- typical and atypical developing children.
- paediatric populations of epilepsy and stroke.
- resting-state and task-based EEG recordings.
- SSVEP and pre-movement/motor-initiation tasks.

Additionally, network analysis was explored from the complementary methodology point-of-view with respect to multi-way structural analysis.

In Chapter 3, a survey of contemporary literature exploring EEG-based BCI in children was presented. The review of literature served to frame the technical work developed in this thesis within the context of the growing field of BCI research. The exploration of literature provided both motivation for improving characterization of paediatric EEG data as well as support for how advancements in this research could translate to important outcomes for children. The chapter also included a look at how the field of EEG-based BCI for children has developed since the initial survey. The findings indicated a growing interest in the field of BCI for children, with the caveat that few technical advancements in the field were focused on improving the characterization of paediatric EEG. This establishes the developments in

this thesis as groundwork for furthering the signal processing aspect of EEG-based BCI for children.

In Chapter 4, and the affiliated Appendix B, the foundation of the proposed multi-way analysis framework was introduced on resting-state EEG data and established across various paediatric datasets and conditions. Building the multi-way model as a constrained 3-way CP model with $[Spatial] \times [Spectral] \times [Subject]$ modes provided a description of the relationships in the rEEG data related to child age, and by proxy, development. A novel component selection step was introduced which exploited the properties of unimodality and CORCONDIA to identify the key number of component factors needed to fully describe underlying developmental information in resting-state EEG data. The constructed multi-way architecture successfully characterized development-specific features across children from the rEEG data based on its validation in 4 real-world and one simulated paediatric datasets. The range of scope in the datasets for both age and developmental conditions (e.g. epileptic, paediatric stroke, or typical) demonstrated the general applicability of the developed methods. Characteristic filters derived from the 3-way CP modelled data aided in assessing classification abilities of the multi-way analysis. The classification results revealed significant improvements for predicting a child's age using the multi-way analysis as compared to random selection or single-class classification schemes, providing evidence of the successful identification of key factors corresponding to child age and development. Evaluating t-SNE maps of the data further revealed that the boons from the multi-way analysis depended on the proper ordering and structure of the 3-mode CP model.

In Chapter 5, the thesis turned toward extending the multi-way framework to the event-triggered, task-based EEG paradigm of the SSVEP, adapted from the Contrast Change Paradigm task. The chapter introduced orthogonal and non-negativity constraints to an extended 4-way CP tensor modelling the SSVEP-driven EEG data, in order to isolate the separate SSVEP signals from each other and noise. The presented work altered the direct projection and characteristic filter methodology to directly assess unseen subject and SSVEP trial data at the level of individual left/right selection trials, as expected in a real-world BCI setting. The developed characteristic filters provided rapid, binary classification of the SSVEP signals in a system which did not need subject-specific classifier calibration. This work was

validated in a BCI setting where SSVEP-driven EEG was used to assess attention of children between left and right objects on a screen, with resulting accuracies of the multi-way model on par with current SSVEP-based BCI paradigms ($\approx 70\%$).

Chapter 6 introduced the novel JEDI model to better encapsulate a wider understanding of child development within the proposed multi-way framework. The JEDI model represented the tensor-matrix structured data fusion of a multi-way CP tensor of an active EEG task and a matrix of clinical measures determining child development via standard psychometric assessments. For the multi-way aspect of the JEDI model, a non-negative 4-way CP tensor which modelled activity in the motor cortex of children during both motor planning (e.g. pre-movement) and actual left/right upper extremity movement (e.g. motor initiation) for a common EEG task (button press) was given. Additionally, the chapter presented a multi-step grid search analysis which determined the hyperparameters weighting the relative tensor, matrix and regularization relationships within the JEDI model's structured data fusion.

The JEDI model was evaluated using a combination of three repeated measures of EEG selection tasks, where two measures were used in determining the JEDI model and its hyperparameter validation, while the third was used as unseen data for testing. Results demonstrated the JEDI model was capable of inferring developmental scores measured by the clinical psychometric tests, with estimated scores falling approximately within the same or adjacent categorical group (e.g., normal, mild impairment, etc.) as the actual clinical values. At the population level, the JEDI model did not significantly differ in its predicted distribution of developmental scores when evaluated against the actual distribution of the clinical measures. The JEDI model offers a novel assessment tool which could pave the way for further development of data-driven methodologies that characterize child development rapidly in clinical settings.

Finally, Chapter 7 explored a complementary developmental analysis scheme to the multi-way framework in the form of network analysis. The chapter established a set of network coupling measures (ICOH, PSI, WPLI) and unbiased graph sub-networks (MST, CST) which interrogated rEEG of very young children with epilepsy to determine correlations to cognitive impairment. Significant correlations to impairment and their associated physiological properties were discussed, with a subset of the correlations chosen via cross-validated

feature selection for classifier training and prediction of cognitive impairment. A clinically relevant cost matrix was developed for the classification of the children into categories of normal development, mild/moderate impairment and severe impairment, weighted to emphasize clinically important features. Results found the network analysis improved on both classification accuracy and reducing misclassification penalties from the cost matrix as compared to random selection and single-class classification schemes, establishing the pertinence of the proposed network analysis. Correlation to the associated findings in the multi-way analysis solidified the complementary nature of the network analysis methodology for both information gained from the rEEG and how to approach characterizing paediatric EEG data.

8.1.1 Positioning within Contemporary Multi-way Research

The contributions developed in this thesis are well positioned within contemporary multi-way analysis research. Multi-way analysis, and in particular its use in conjunction with biomedical data, is an emerging field for scientists around the globe, which has yet to be fully explored [26]. Research in multi-way analysis spans from improving the algorithms underlying multi-way model decomposition (e.g. incorporating machine learning techniques [210–213]) to extending the feature extraction capabilities [149, 170, 214, 215].

Contributions from the thesis fall under the particular umbrella of investigating how to integrate multi-way analysis structures to specific biomedical signals (i.e. paediatric EEG) in order to elucidate information of interest. Motivation for the thesis work originated with advancing the signal processing behind BCI technology for children. This motivation is reflected by other researchers in the multi-way analysis field, as seen in the recent state-of-the-art review of the classification algorithms for EEG-based BCI given by Lotte *et al.* [26]. In their work, Lotte *et al.* highlight feature extraction and classification via multi-way tensor analysis as an emerging and promising tool for analysis of EEG data, while noting its relative infancy with respect to established research and development in EEG applications [26]. The work and contributions of this thesis thus helps contribute to better establishing the research foundation of multi-way methodology in EEG characterization, which is highlighted by Lotte and company as missing. The contributions in Chapters 4-6 provide example blue-prints of higher-order CP models of paediatric EEG for a growing number of domains (e.g. 3-way, 4-way, 5-way), supplementing

other developments in representing adult EEG data as multi-way arrays [38, 150, 215, 216].

Similarly, the methodology in Chapter 5 to characterize SSVEP-driven EEG signals can be considered as a companion research piece to the work done by Zink *et al.* [150]. Where the presented work focused on a binary visual stimulus (e.g., left/right SSVEP selection task), Zink and associates developed the multi-way framework for a binary auditory stimulus (e.g., two-tone auditory selection task) [150]. For their multi-way framework, Zink explored both a CP model and a block-term decomposition (BTD) model [37, 157, 217] for their study, finding the BTD was capable of better performance than the CP model [150]. It should be noted, however, that no constraints were used with the CP model in [150], thereby potentially affecting its usefulness in representing the EEG data [40, 143, 150, 217]. The work in Chapter 5 thus contributes an alternative view of the CP tensor framework compared to [150], emphasizing how appropriate constraints can play a particularly important role which improves effective modelling of the data. Developments in Chapter 5, in conjunction with findings from Zink, help derive a set of methodologies suitable for simple EEG selection tasks, where dominating signals can be easily identified using a multi-way framework. These developments may be incorporated with other modern research into SSVEP analysis for children, like the work in [135], to help further cement the foundation of multi-way analysis in EEG-based applications.

Other current advancements in the niche of multi-way analysis research in biomedical signals focuses on exploiting tensor-matrix joint factorization of signals [218–220], similar in spirit to the developments of the JEDI model. The focus for these studies, however, is often the joint factorization of EEG data with another biomedical signal recording like functional magnetic resonance imaging (fMRI), as seen in [218–220], or magnetoencephalography (MEG), as seen in [221]. Developments from the JEDI model in Chapter 6 thus contribute to this subset of modern multi-way research by establishing an alternative tensor-matrix structured data fusion scheme, where the jointly factorized matrix represents a wholly separate type of measurement, e.g. psychometric data.

These examples highlight how the contributions of this thesis are well integrated in the current state of multi-way analysis research.

8.2 Limitations

Some limitations occurring in this thesis are related to the sample sizes of the real paediatric EEG datasets used. Often, the population size is under 50 for the total number of children assessed, with the actual children per age-group proportionally smaller, usually between 1 and 14. Larger populations would allow for more powerful statistical testing, more robust cross-validation and greater representation of the cross-sectional paediatric population for more convincing results. This limitation is in no way unique to this thesis as obtaining EEG and other data from large populations is difficult, especially for paediatric and/or pathological patients. To help mitigate this, this thesis has used multiple datasets, simulation data and appropriate repeated recordings (e.g. the non-longitudinal repeated task recordings of the CMI dataset) where possible to provide more conclusive evidence that the results are representative across different sections of the paediatric population.

Similarly, the paediatric EEG datasets evaluated in this thesis were not recorded under uniform conditions, locations or by the same research team due to the limited availability of paediatric EEG data. By utilizing both collaborations with hospitals and publicly accessible datasets to establish the proposed methods, the results represent a range of relevant recordings for analysis. Being able to recruit children for specifically designed EEG-based paradigms with the EEG recorded under homogeneous settings would help alleviate potential systemic errors from these variable recording conditions. However, the variation in the recording conditions do provide support for the general applicability of the developed methods, demonstrating their relevance across a range of settings. Nonetheless, having homogeneous, specific EEG-tasks of interest would help bolster the translational relevance of the multi-way research with respect to EEG-driven applications like BCI.

Additionally, nearly all the paediatric EEG data examined in these studies lacked a longitudinal component. Particularly, following up with the children months or years after initial developmental assessment and EEG recordings would help reinforce the multi-way framework's ability to track development-related features. Incorporating in a subset of the typical and atypical developing patients who received intervention of some sort could help establish if the multi-way analysis could also distinguish between developmental changes

associated with ageing and those associated with the active intervention. Ideally, in the future, such a dataset will become available and the investigations into the utility of actively assessing these various developmental patterns via the multi-way framework can be further evaluated.

The choice to focus on the CP model throughout the thesis brought on some restrictions as well, in terms of its inflexible nature. The strict superdiagonal core of the CP model restricts the ability to fit a variable number of components across different domains, which might in turn better model the underlying EEG data structures. In other words, the CP model's superdiagonal core only permits limited, specific interaction between the factor components of different modes at the unique locations on the CP models core, e.g. the diagonal seen in Figure 2.8). All other interaction between the components is not considered. In more flexible cores this is not true, e.g. the various interactions seen off-diagonal in Figure 2.9 representing a non-superdiagonal core. The trade-off, of course, is that the CP model gave component factors guaranteed to interact in a predictable manner, allowing us to interpret and understand the components' associations across the different domains. This trade-off was justifiable, as the present work focused on laying the multi-way framework foundation.

8.3 Future Work

There are many avenues for future work which could expand upon the contributions of this thesis, due to its broad applicability in the characterization of EEG data. Notably, future work could include exploring alternative tensor models with more flexibility, like PARAFAC2 [39, 50, 222], BTD [37, 157, 217] or Tucker [43, 215, 223]. PARAFAC2 is a good option for implementing some flexibility into the tensor framework presented, as it largely retains the benefits of the CP model but offers a little more flexibility through permitting variation in components for a single domain [39, 50, 222]. As such, PARAFAC2 represents a tensor modelling option which resembles the CP model with slightly more flexibility, which can lead to benefits in biomedical signal investigations [39, 50]. Similarly, BTD models combine the CP and Tucker models to create an 'intermediate' flexible model, with advantages gained from both [37, 157, 217]. The BTD tensor model represents higher-order data using a sum of Tucker models constrained to include a set of factor matrices with specific properties such as full column rank and a sub-tensor decomposed by higher-order SVD [37, 157, 217]. More

flexible tensor models like these could be useful in improving model fitting in the EEG data, and investigations could include how to balance the model's flexibility and output interpretability.

Another avenue for further investigation is the translation of the proposed techniques from off-line analysis post recording to online analysis, done in a real-time BCI setting like in [224]. The characteristic filters could potentially be first identified and calculated off-line from previous EEG data recordings, and then implemented in the real-time online BCI for classification. Furthermore, consideration should be given to how the characteristic filters could be updated as the child develops. This may require introducing a new mode in the data, where unique sessions are recorded and compared to see assess changes in the derived characteristic filters over time. Additionally, concerns regarding recalibration of the filters and their role in assessing development would need to be addressed. In particular, this line of investigations is plausible for extending the SSVEP-based contributions, where the online classification could rapidly be discerned by the characteristic filters. Establishing this type of scenario is an active ongoing collaboration currently in its infancy with the researchers in Norton *et al.* [31, 135].

The success of the JEDI model and the network analysis highlighted how developmental information may be encoded in multiple ways within paediatric EEG data. Unveiling how this encoded developmental information may be shared and exploited between the JEDI model and the network analysis approaches is another aspect to consider for future work. Integrating in results from the network analysis to a structured data fusion with the JEDI model, or some variant on it with a more flexible tensor core, could help reveal important ties between the multi-way framework and the network analysis methodology. This could also include how to incorporate the previous tensor decomposition/network analysis results (e.g. from a prior session) into the present joint factorization, similar to research in 'supervised' tensor factorizations [225–227]. Furthermore, adding additional measures of development outside of cognition is another line of future work to be investigated by both the network analysis and the JEDI model, or even potentially by some combined data fusion model between the two. The contributions established in the thesis help provide the foundations needed to investigate further development of these EEG developmental characterization tools.

8.4 Conclusions

Overall, developments of this thesis examine a novel means for characterizing paediatric EEG using multi-way analysis, demonstrating its potential in a swathe of child datasets comprising assorted conditions, age ranges and developmental status. The novel investigations established the versatility of the multi-way framework in analyzing paediatric EEG and provided a suitable foundation for continued development of this research topic. It is expected that the impact of the developed methods will extend beyond paediatric EEG analysis, to the wider field of BCI research, as well as other bioengineering and biomedical problems which could benefit from delineating relationships in higher-order data. May the analytical safari provided by this thesis inspire others towards creative, scientific thinking in their own promotion and development of translational research.

Appendix A: List of Publications

List of Accepted Journal Publications

1. **Kinney-Lang, E.**, Auyeung, B. and Escudero, J., 2016. Expanding the (kaleido) scope: exploring current literature trends for translating electroencephalography (EEG) based brain–computer interfaces for motor rehabilitation in children. *Journal of neural engineering*, 13(6), p.061002 [1].
2. **Kinney-Lang, E.**, Spyrou, L., Ebied, A., Chin, R.F. and Escudero, J., 2018. Tensor-driven extraction of developmental features from varying paediatric EEG datasets. *Journal of neural engineering*, 15(4), p.046024 [3].
3. **Kinney-Lang, E.**, Ebied, A., Auyeung, B., Chin, R.F., and Escudero, J., 2018. Introducing the Joint EEG-Development Inference (JEDI) Model: A Multi-way, Data Fusion Approach for Estimating Paediatric Developmental Scores Via EEG. *IEEE Transactions on neural systems & rehabilitation engineering*, [5].
4. **Kinney-Lang, E.**, Yoong, M., Hunter, M., Tallur, K.K., Shetty, J., McLellan, A., Chin, R.F. and Escudero, J., 2017. Analysis of EEG networks and their correlation with cognitive impairment in preschool children with epilepsy. *Epilepsy & Behaviour*, 90, p.45-56 [6].

List of Accepted Conference Publications

1. **Kinney-Lang, E.**, Spyrou, L., Ebied, A., Chin, R. and Escudero, J., 2017, July. Elucidating age-specific patterns from background electroencephalogram pediatric datasets via PARAFAC. *In Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE* (pp. 3797-3800). IEEE [2].
2. **Kinney-Lang, E.**, Ebied, A., and Escudero, J., 2018. Building a Tensor Framework for the Analysis and Classification of Steady-State Visual Evoked Potentials in Children. *Eusipco'18: Proceedings of the 26th European Signal Processing Conference*. ISBN 978-90-827970-1-5. *EURASIP 2018*, [4].

Appendix B: Evaluating rEEG in Children with Paediatric Stroke

This appendix builds upon the framework presented in Chapter 4, offering a pilot study of using multi-way analysis to characterize EEG recordings of a paediatirc stroke population. The study aimed to replicate the findings from Chapter 4 in a population who could readily benefit from BCI implementation.

8.5 Materials

8.5.1 Alberta Children's Hospital Dataset

A small cohort of resting-state EEG data from 16 pre-adolescent children aged 1-11 years old was provided by the Alberta Children's Hospital, henceforth referred to as the ACH dataset. Children in the ACH dataset all suffered from an early-life arterial ischemic stroke (AIS), leading to localized lesions on the cortex. The resting-state data was collected as part of their clinical evaluation, resulting in a high likelihood that the selected subjects may suffer from additional complications like epilepsy. Two subjects had significantly fewer EEG channels than the rest and were therefore excluded from the study, resulting in $n = 14$ subjects for analysis. Due to strict privacy policies, all pre-processing on the data had to be done by the Alberta Children's Hospital staff, with only the Fourier-transformed EEG spectra given for analysis. As such, we requested for all children to approximately follow the pre-processing used for the MMEC and CHB-MIT datasets, outlined in Section 4.3. We were assured all potential contaminations of the EEG by artifacts, including potential epileptic activity, were excluded in the provided data.

Importantly, one of the ACH subjects had a longitudinal set of EEG recordings. Their EEG was recorded at both one and five years of age. Although this was only a single child, we investigated the longitudinal data as a separate tensor to see if the proposed analysis could capture expected developmental features at these separate time points for the same child.

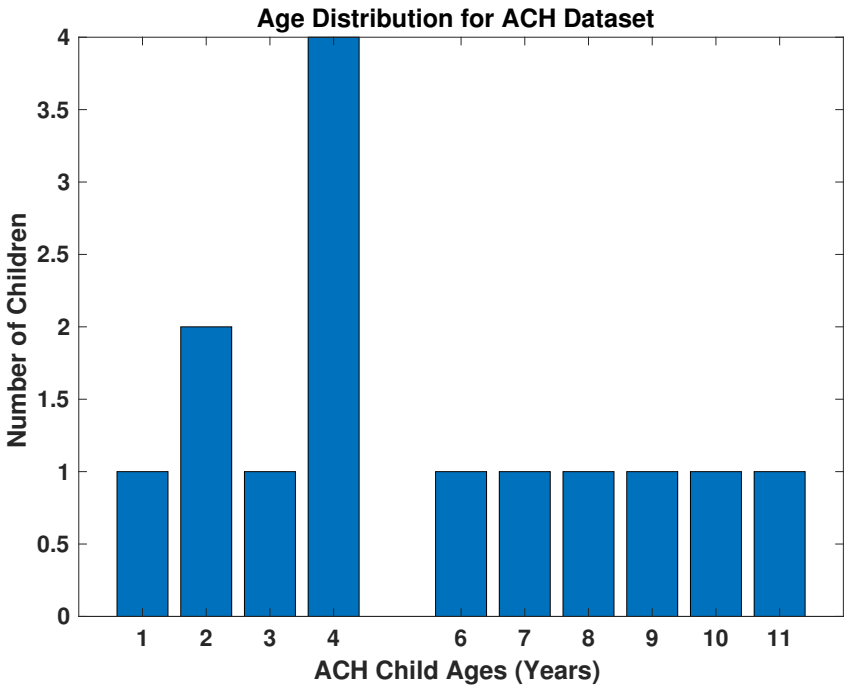

Figure 8.1: Distribution of subjects per age for the ACH dataset.

Again, the pre-processing was done by the ACH, and only their Fourier-transformed EEG spectra was provided to us for analysis.

A small illustration of the distribution of ages in the ACH data is seen in Figure 8.1.

8.6 Methods

EEG analysis replicated the methods described in Chapter 4. Of note, the ACH removed artifacts using the Matlab plugin Automagic, similar to the pre-processed CMI data described in Chapter 5. Additionally, frequency resolution was given to be 0.2 Hz over the approximate [0.5-30] Hz band of interest for the ACH data. From this, a three-way tensor consisting of $[Spatial] \times [Spectral] \times [Subject]$ modes was created using the EEG channel, power spectra and subject age data for the ACH dataset, resulting in $(19) \times (148) \times (14)$ elements. For the longitudinal data, the three-way tensor had $(19) \times (148) \times (2)$ elements, with the $[Subject]$ domain representing the two longitudinal recording points.

8.7 Results

Results from the ACH dataset provide a preliminary exploration of the proposed tensor framework for children who would readily benefit from BCI applications, e.g. have impaired motor functioning due to paediatric stroke. Similar to the CHB-MIT dataset, the small pilot data from the ACH does not include enough children at each age to effectively evaluate the data using classification (see Figure 8.1 for age distributions in the ACH data). However, examining the resulting profiles given by the CP decomposition can provide valuable insight into how the proposed CP analysis can capture developmental features. Thus, the resulting successful identificaiton of feature profiles from the CP model decomposition are presented in Figure 8.2. The CP model decomposition for the ACH had a CORCONDIA slightly below the target of 70%, at 65.5%.

8.7.1 Longitudinal Data

Included in the ACH data was a single subject who was recorded twice over the span of 4 years. The subject originally had EEG recording done after the initial paediatric stroke at 1-year-old, and had EEG recordings done again at 5-years of age. Although this is only a single patient, the two longitudinal recordings provide a snapshot into whether the CP tensor analysis can identify developmental progress in a single child's EEG over time. The tensor in this situation is built using the $[Spatial]$ and $[Spectral]$ modes similar to Figure 8.2, but with the $[Subject]$ domain containing only their two recordings at age one and five. The resulting feature profiles identified by the CP decomposition are illustrated in Figure 8.3. The $[Spatial]$, $[Spectral]$ and $[Subject]$ domains are represented in the 3 panels from top to bottom, respectively. The CP model of the longitudinal data had a CORCONDIA value of 85.7%, and with nearly 98% explained variance.

8.8 Discussion

8.8.1 Developmental Feature Profiles from the Paediatric Stroke Data

The shifting developmental pattern is again picked up for both the standard and longitudinal factor profiles from the CP analysis of the ACH data. The prominent factor profiles seen in

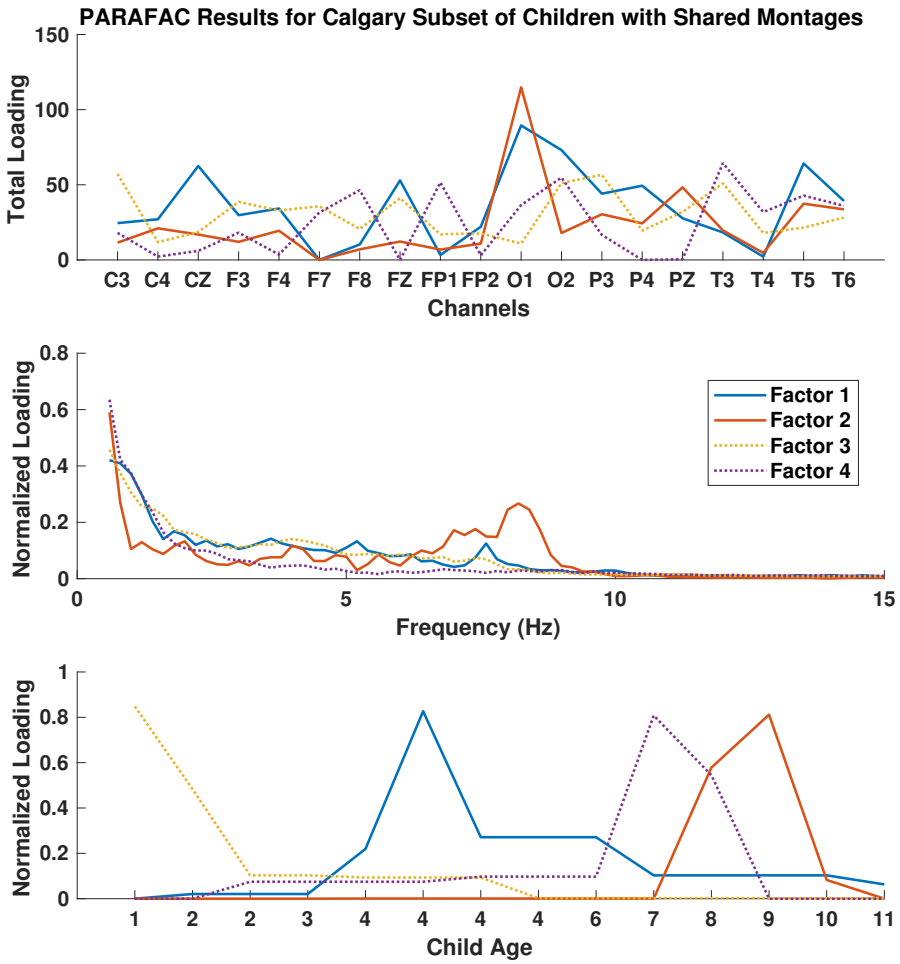

Figure 8.2: The graphical representation of the feature profiles identified from the ACH dataset for paediatric stroke patients from age 1-11 using the CP decomposition. The results are presented as a 3 panel set of graphs, corresponding to the $[Spatial]$, $[Spectral]$ and $[Subject]$ domains from top to bottom, respectively. Factor 1 and 2 capture expected developmental differences across the ages, while factor 3 and 4 appear to effectively capture differences in background noise during development.

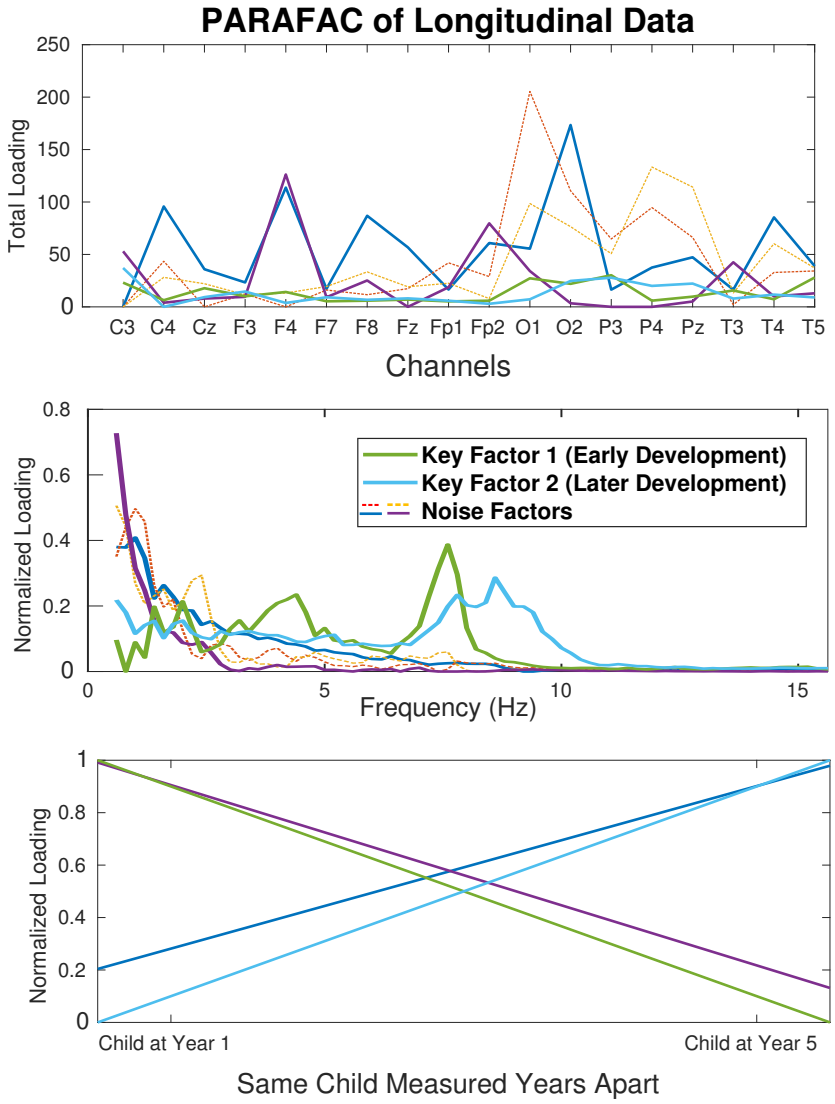

Figure 8.3: A representation of the feature profiles identified by the CP tensor decomposition, in longitudinal EEG recordings at one and five years old for a single patient. The results are presented as a 3 panel set of graphs, corresponding to the $[Spatial]$, $[Spectral]$ and $[Subject]$ domains from top to bottom, respectively, with the $[Subject]$ domain consisting of the two recording time-points. A low value (near 0) in the $[Subject]$ domain loading indicates little contribution from that recording time point, while a high value (near 1) reflects a significant contribution.

Figure 8.2 appear to successfully match developmental trends from the literature and other datasets [12, 21, 22, 28]. The older $[Subject]$ domain profiles correspond to a shift towards more dominating frequencies in the $[Spectral]$ domain near the expected alpha frequency range compared to younger subjects. This shift is likely represented by the differences in peaks at 7 Hz and 8 Hz for factor 1 and 2, reflecting changing structures for the younger and older children respectively. While this difference in peaks is not as pronounced as the other datasets, e.g. the MMEC in Chapter 4, its result is not necessarily surprising. The small size of the ACH dataset along with the extremely heterogeneous nature of AIS in children could have limited the effectiveness of the proposed CP model. This discrepancy could potentially be alleviated in task-oriented EEG recordings which would relax our unimodal constraint, and by incorporating a more complex representation for patient development besides age, to account for how the paediatric stroke may have affected the child's development.

Similar to the findings in the CHB-MIT and MMEC data from Chapter 4, factor 3 and 4 appear to also have captured innate differences in background noise for the younger and older children, respectively. Factor 3 dominates the low-frequency activity of the brain in the $[Spectral]$ domain consistently over factor 4, which considered alongside their $[Subject]$ domain profiles, aligns with the expected behaviour in EEG power changing during child development [12, 21, 28]. Therefore the CP model was able to successfully identify two important developmental features for the ACH dataset in factors 1 and 2, while effectively separating background noise between younger and older children in factors 3 and 4.

Evaluating the limited longitudinal ACH data reveals encouraging results. The feature profiles in Figure 8.3 help verify that the CP model can effectively capture developmental profiles of interest in the same child over time, complementing the other cross-sectional analysis results. The key factor profiles identified from the CP decomposition reflect the expected developmental behaviour just like the cross-sectional results, with the earlier recording having a higher-power peak lower in frequency, which then shifts to a lower-power peak in higher frequencies in the later recording [12, 21, 28]. Again, the CP model is able to distinguish factors representative of unique developmental background noise between the two time-point recordings, as seen in the cross-sectional MMEC, CHB-MIT and other datasets. The $[Subject]$ domain captures how the distinct feature profiles at the two longitudinal recordings interact in

the $[Spectral]$ and $[Spatial]$ domains. The low values in the $[Subject]$ domain, e.g. values near zero, indicate little contribution by the factors at that recording time-point while high values, e.g. near one, indicate significant contribution. Together, these replicated results on the limited longitudinal data suggest it could be possible to track development of a child as they age using the CP tensor framework to assess their changing development via EEG. This, however, would require more nuance in terms of what is defined as child development, besides the proxy of chronological age. This topic was explored in part in Chapter 6, through the use of the Joint EEG Development Inference (JEDI) model.

8.8.2 Limitations

The results from this pilot study should be considered carefully due to some limitations in the dataset. The number of children included was relatively small, and only had a single or few subjects per age. Additionally, EEG recordings in young children are typical for identifying if a child has epilepsy. In this manner, the dataset skews towards a subset of the paediatric stroke population who potentially have epilepsy co-occurring with the AIS. Finally, the deficits resulting from the AIS are likely quite heterogeneous in their presentation, due to differences in trauma such as location and size of the brain lesion. To address these limitations, a much larger, robust dataset of children with early-life AIS should be investigated, with sub-categories for different lesions considered.

Appendix C: Funding Support Acknowledgements

Funding for this thesis was provided, in part, by several different sources. This appendix aims to acknowledge specific sources of funding and the associated institutions which supported the research presented.

For the data provided by the Muir Maxwell Epilepsy Centre (Chapter 4, 7), the author would like to thank the patients and families who participated in the NEUROPROFILES [140] study, and the clinical team who provided the data. In particular, Matthew Hunter, Michael Yoong, Jay Shetty, Krishnaraya Kamath Tallur, Ailsa McLellan and Richard Chin. Funding for the research presented in Chapter 4, and Chapter 7 using the MMEC data was provided by the RS McDonald Trust, Thomas Theodore Scott Ingram Memorial Fund, and the Muir Maxwell Trust.

For the data provided by the Alberta Children's Hospital (Appendix B), the author would like to thank Edinburgh Neuroscience, for providing financial support in establishing a collaboration via the Edinburgh NeuroResearcher's Grant. Additionally, the author thanks Dr. Kirton's Lab at the University of Calgary and the Alberta Children's Hospital for support in the collaboration.

Bibliography

[1] E. Kinney-Lang, B. Auyeung, and J. Escudero, "Expanding the (kaleido)scope: exploring current literature trends for translating electroencephalography based brain–computer interfaces for motor rehabilitation in children," *J. Neural Eng.*, vol. 13, no. 6, p. 061002, 2016. [Online]. Available: http://stacks.iop.org/1741-2552/13/i=6/a=061002?key=crossref.1a3945e388543cb30cfd62d3565ed379

[2] E. Kinney-Lang, L. Spyrou, A. Ebied, R. Chin, and J. Escudero, "Elucidating age-specific patterns from background electroencephalogram pediatric datasets via PARAFAC," in *2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* IEEE, jul 2017, pp. 3797–3800. [Online]. Available: http://ieeexplore.ieee.org/document/8037684/

[3] E. Kinney-Lang, L. Spyrou, A. Ebied, R. F. M. Chin, and J. Escudero, "Tensor-driven extraction of developmental features from varying paediatric EEG datasets," *J. Neural Eng.*, vol. 15, no. 4, p. 046024, aug 2018. [Online]. Available: http://iopscience.iop.org/article/10.1088/1741-2552/aac664

[4] E. Kinney-Lang, A. Ebied, and J. Escudero, "Building a Tensor Framework for the Analysis and Classification of Steady-State Visual Evoked Potentials in Children," in *2018 26th Eur. Signal Process. Conf.* IEEE, sep 2018, pp. 296–300. [Online]. Available: https://ieeexplore.ieee.org/document/8553012/

[5] E. Kinney-lang, A. Ebied, B. Auyeung, R. Chin, and J. Escudero, "Introducing the joint eeg-development inference (jedi) model: A multi-way, data fusion approach for estimating paediatric developmental scores via eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 1 2019.

[6] E. Kinney-Lang, M. Yoong, M. Hunter, K. Kamath Tallur, J. Shetty, A. McLellan, R. FM Chin, and J. Escudero, "Analysis of EEG networks and their correlation with cognitive impairment in preschool children with epilepsy," *Epilepsy Behav.*, vol. 90, pp. 45–56, jan 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1525505018306784

[7] A. Wachoski and L. Wachoski, "The matrix-film," 1999.

[8] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002. [Online]. Available: www.elsevier.com/locate/clinphhttp://www.ncbi.nlm.nih.gov/pubmed/12048038

[9] S. Kotermanski and M. Cascio, "Neuronal Action Potentials and Ion Channel Allostery," in *Encycl. Cell Biol.* Elsevier, 2016, vol. 1, pp. 244–251. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/B9780123944474100288

[10] E. K. S. Louis and L. C. Frey, *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*, 1st ed., E. K. St. Louis and L. C. Frey, Eds. Chicago, IL: American Epilepsy Society, 2016. [Online]. Available: https://www.osapublishing.org/abstract.cfm?URI=boe-7-7-2524

[11] F. Lopes da Silva, "EEG and MEG: Relevance to neuroscience," pp. 1112–1128, 2013.

[12] V. Miskovic, X. Ma, C.-A. Chou, M. Fan, M. Owens, H. Sayama, and B. E. Gibb, "Developmental changes in spontaneous electrocortical activity and network organization from early to late childhood." *Neuroimage*, vol. 118, pp. 237–47, sep 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1053811915005108

[13] M. Zaepffel, R. Trachel, B. E. Kilavik, and T. Brochier, "Modulations of EEG Beta Power during Planning and Execution of Grasping Movements," *PLoS One*, vol. 8, no. 3, p. e60060, mar 2013. [Online]. Available: http://dx.plos.org/10.1371/journal.pone.0060060

[14] F. Darvas, R. Scherer, J. G. Ojemann, R. P. Rao, K. J. Miller, and L. B. Sorensen, "High gamma mapping using EEG," *Neuroimage*, vol. 49, no. 1, pp. 930–938, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2009.08.041

[15] H. Song, D. Zhang, Z. Ling, H. Zuo, and B. Hong, "High gamma oscillations enhance the subdural visual speller," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 1711–1714, 2012.

[16] G. Pfurtscheller and F. H. Lopes, "Event-related EEG / MEG synchronization and desynchronization : basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842–1857, 1999.

[17] J. Escudero, E. Ifeachor, J. P. Zajicek, C. Green, J. Shearer, and S. Pearson, "Machine learning-based method for personalized and cost-effective detection of Alzheimer's disease." *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 164–8, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22893371

[18] I. Daly, R. Scherer, M. Billinger, and G. Müller-Putz, "FORCe: Fully Online and automated artifact Removal for brain-Computer interfacing." *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 4320, no. c, pp. 1–13, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25134085

[19] R. Scherer, M. Billinger, J. Wagner, A. Schwarz, D. T. Hettich, E. Bolinger, M. Lloria Garcia, J. Navarro, and G. Müller-Putz, "Thought-based row-column scanning communication board for individuals with cerebral palsy," *Ann. Phys. Rehabil. Med.*, vol. 58, no. 1, pp. 14–22, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25661447

[20] B. Boashash, N. Ali, T. Ben-jabeur, N. A. Khan, and T. Ben-jabeur, "Time – frequency features for pattern recognition using high-resolution TFDs : A tutorial review," *Digit. Signal Process.*, vol. 40, no. 1, pp. 1–30, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.dsp.2014.12.015

[21] M. Matsuura, K. Yamamoto, H. Fukuzawa, Y. Okubo, H. Uesugi, M. Moriiwa, T. Kojima, and Y. Shimazono, "Age development and sex differences of various EEG elements in healthy children and adults–quantification by a computerized wave form recognition method," *Electroencephalogr Clin Neurophysiol*, vol. 60, no. 5, pp. 394–406, 1985. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/2580690

[22] P. J. Marshall, Y. Bar-Haim, and N. A. Fox, "Development of the EEG from 5 months to 4 years of age," *Clin. Neurophysiol.*, vol. 113, no. 8, pp. 1199–1208, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1388245702001633

[23] H. Schmidt, W. Woldman, M. Goodfellow, F. A. Chowdhury, M. Koutroumanidis, S. Jewell, M. P. Richardson, and J. R. Terry, "A computational biomarker of idiopathic generalized epilepsy from resting state EEG," *Epilepsia*, vol. 57, no. 10, pp. e200–e204, 2016.

[24] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," pp. 2266–2301, oct 2005. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245705002403

[25] N. Langer, E. J. Ho, L. M. Alexander, H. Y. Xu, R. K. Jozanovic, S. Henin, A. Petroni, S. Cohen, E. T. Marcelle, L. C. Parra, M. P. Milham, and S. P. Kelly, "A resource for assessing information processing in the developing brain using EEG and eye tracking," *Sci. Data*, vol. 4, p. 170040, apr 2017. [Online]. Available: http://www.nature.com/articles/sdata201740

[26] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," p. 031005, jun 2018. [Online]. Available: http://stacks.iop.org/1741-2552/15/i=3/a=031005?key=crossref.9cd2b15ab65c8ad34b475584b43dc509

[27] B. Pletzer, H. Kerschbaum, and W. Klimesch, "When frequencies never synchronize: The golden mean and the resting EEG," *Brain Res.*, vol. 1335, pp. 91–102, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.brainres.2010.03.074

[28] T. Gasser, R. Verleger, P. Bächer, and L. Sroka, "Development of the EEG of school-age children and adolescents. I. Analysis of band power," *Electroencephalogr.*

Clin. Neurophysiol., vol. 69, no. 2, pp. 91–99, feb 1988. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0013469488902040

[29] R. Srinivasan, "Spatial structure of the human alpha rhythm: global correlation in adults and local correlation in children." *Clin. Neurophysiol.*, vol. 110, no. 8, pp. 1351–62, 1999. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10454270

[30] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review." *Sensors (Basel).*, vol. 12, no. 2, pp. 1211–79, 2012. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/22438708

[31] J. L. Mullins, "SSVEP-BASED BCI PERFORMANCE IN CHILDREN," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2017. [Online]. Available: https://www.ideals.illinois.edu/bitstream/handle/2142/88086/MULLINS-THESIS-2015.pdf?sequence=1{&}isAllowed=y

[32] M. J. Taylor and T. Baldeweg, "Application of EEG, ERP and intracranial recordings to the investigation of cognitive functions in children," *Dev. Sci.*, vol. 5, no. 3, pp. 318–334, 2002.

[33] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations*. Chichester, UK: John Wiley & Sons, Ltd, sep 2009. [Online]. Available: http://doi.wiley.com/10.1002/9780470747278

[34] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, vol. 374, no. 2065, p. 20150202, 2016. [Online]. Available: http://rsta.royalsocietypublishing.org/lookup/doi/10.1098/rsta.2015.0202

[35] A. Ebied, E. Kinney-Lang, L. Spyrou, and J. Escudero, "Evaluation of matrix factorisation approaches for muscle synergy extraction," *Med. Eng. Phys.*, vol. 57, pp. 51–60, jul 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1350453318300572

[36] M. C. Tresch and V. C. K Cheung, "Matrix Factorization Algorithms for the Identification of Muscle Synergies: Evaluation on Simulated and Experimental Data Sets," *J Neuro-physiol*, vol. 95, pp. 2199–2212, 2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16394079

[37] A. Cichocki, "Tensor Decompositions: A New Concept in Brain Data Analysis?" *arXiv Prepr. arXiv1305.0395*, vol. 50, no. 2011, pp. 507–517, may 2013. [Online]. Available: http://arxiv.org/abs/1305.0395

[38] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, 2015.

[39] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 1, pp. 6–20, 2009.

[40] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," *Nonlinear Theory Its Appl. IEICE*, vol. 1, no. 1, pp. 37–68, 2010. [Online]. Available: http://joi.jlc.jst.go.jp/JST.JSTAGE/nolta/1.37?from=CrossRef

[41] A. Cichocki, Y. Washizawa, T. Rutkowski, H. Bakardjian, Anh-Huy Phan, Seungjin Choi, Hyekyoung Lee, Qibin Zhao, Liqing Zhang, and Yuanqing Li, "Noninvasive BCIs: Multiway Signal-Processing Array Decompositions," *Computer (Long. Beach. Calif).*, vol. 41, no. 10, pp. 34–42, oct 2008. [Online]. Available: http://ieeexplore.ieee.org/document/4640658/

[42] T. G. Kolda and B. W. Bader, "Tensor Decompositions and Applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2008. [Online]. Available: http://epubs.siam.org/doi/abs/10.1137/07070111X

[43] F. Cong, Q. H. Lin, L. D. Kuang, X. F. Gong, P. Astikainen, and T. Ristaniemi, "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027015001016

[44] Y. Liu, M. Li, H. Zhang, H. Wang, J. Li, J. Jia, Y. Wu, and L. Zhang, "A tensor-based scheme for stroke patients' motor imagery EEG analysis in BCI-FES rehabilitation training," *J. Neurosci. Methods*, vol. 222, pp. 238–249, 2014.

[45] N. Klaas, M. Faber, R. Bro, and P. K. Hopke, "Recent developments in CANDECOMP / PARAFAC algorithms : a critical review," *Chemom. Intell. Lab. Syst.*, vol. 65, pp. 119 – 137, 2003.

[46] R. a. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis," *UCLA Work. Pap. Phonetics*, vol. 16, no. 10, pp. 1– 84, 1970.

[47] T. G. Kolda and B. W. Bader, "Algorithm 862: MATLAB tensor classes for fast algorithm prototyping," *ACM Trans. Math. Softw*, vol. 32, no. 4, pp. 635–653, 2006. [Online]. Available: https://www.sandia.gov/~tgkolda/pubs/pubfiles/SAND2004-5187.pdf

[48] R. Bro and H. a.L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemom.*, vol. 17, no. 5, pp. 274–286, jun 2003. [Online]. Available: http://doi.wiley.com/10.1002/cem.801

[49] E. E. Papalexakis, "Automatic Unsupervised Tensor Mining with Quality Assessment," in *Proc. 2016 SIAM Int. Conf. Data Min.* Philadelphia, PA: Society for Industrial and Applied Mathematics, jun 2016, pp. 711–719. [Online]. Available: http://epubs.siam.org/doi/pdf/10.1137/1.9781611974348.80

[50] J. Escudero, E. Acar, A. Fernández, and R. Bro, "Multiscale entropy analysis of resting-state magnetoencephalogram with tensor factorisations in Alzheimer's disease," *Brain Res. Bull.*, vol. 119, pp. 136–144, oct 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0361923015000763

[51] N. Lee and A. Cichocki, "Fundamental Tensor Operations for Large-Scale Data Analysis in Tensor Train Formats," *arXiv*, 2014. [Online]. Available: https://arxiv.org/pdf/1405.7786.pdf

[52] A. Cichocki, "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions," *arXiv*, pp. 1–30, mar 2014. [Online]. Available: http://arxiv.org/abs/1403.2048

[53] I. V. Oseledets, "Tensor-Train Decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011. [Online]. Available: http://epubs.siam.org/doi/10.1137/090752286

[54] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, mar 2009. [Online]. Available: http://www.nature.com/articles/nrn2575

[55] C. J. Stam and J. C. Reijneveld, "Graph theoretical analysis of complex networks in the brain," *Nonlinear Biomed. Phys.*, vol. 1, no. 1, p. 3, 2007. [Online]. Available: http://nonlinearbiomedphys.biomedcentral.com/articles/10.1186/1753-4631-1-3

[56] C. J. Stam, "Modern network science of neurological disorders," *Nat. Rev. Neurosci.*, vol. 15, no. 10, pp. 683–695, 2014. [Online]. Available: http://www.nature.com/doifinder/10.1038/nrn3801

[57] M. E. J. Newman, "Mathematics of networks," in *Networks*, 2010, pp. 109–167. [Online]. Available: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199206650.001.0001/acprof-9780199206650-chapter-6

[58] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–9, 2011. [Online]. Available: http://www.hindawi.com/journals/cin/2011/156869/

[59] C. A. Andersson and R. Bro, "The N-way Toolbox for MATLAB," *Chemom. Intell. Lab. Syst.*, vol. 52, no. 1, pp. 1–4, aug 2000. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S016974390000071X

[60] N. Vervliet, O. Debals, and L. De Lathauwer, "Tensorlab 3.0 - Numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization," in *Conf. Rec. - Asilomar Conf. Signals, Syst. Comput.* IEEE, nov 2017, pp. 1733–1738. [Online]. Available: http://ieeexplore.ieee.org/document/7869679/

[61] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," *ACM SIGKDD Explor.*, vol. 11, no. 1, pp. 10–18, 2009. [Online]. Available: https://dl.acm.org/citation.cfm?doid=1656274.1656278

[62] J. R. Wolpaw, "Brain-computer interfaces as new brain output pathways." *J. Physiol.*, vol. 579, no. 3, pp. 613–619, 2007. [Online]. Available: http://doi.wiley.com/10.1113/jphysiol.2006.125948

[63] J. D. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia, "Combining Brain-Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges." *Front. Neurosci.*, vol. 4, no. September, pp. 1–15, 2010. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnins.2010.00161/abstract

[64] H. Cecotti, "Spelling with non-invasive Brain-Computer Interfaces–current and future trends." *J. Physiol. Paris*, vol. 105, no. 1-3, pp. 106–14, 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/21911058

[65] A. Mora-Cortes, N. V. Manyakov, N. Chumerin, and M. M. Van Hulle, "Language Model Applications to Spelling with Brain-Computer Interfaces," *Sensors*, vol. 14, pp. 5967–5993, 2014. [Online]. Available: https://www.mdpi.com/1424-8220/14/4/5967

[66] J. J. Daly and J. R. Wolpaw, "Brain–computer interfaces in neurological rehabilitation," *Lancet Neurol.*, vol. 7, no. 11, pp. 1032–1043, nov 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1474442208702230

[67] L. van Dokkum, T. Ward, and I. Laffont, "Brain computer interfaces for neurorehabilitation – its current status as a rehabilitation strategy post-stroke," *Ann. Phys. Rehabil. Med.*, vol. 58, no. 1, pp. 3–8, feb 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877065714018338

[68] E. V. C. Friedrich, N. Suttie, A. Sivanathan, T. Lim, S. Louchart, and J. A. Pineda, "Brain-computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum." *Front. Neuroeng.*, vol. 7, no. July, p. 21, jul 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25071545

[69] A. Zuberer, D. Brandeis, and R. Drechsler, "Are treatment effects of neurofeedback training in children with ADHD related to the successful regulation of brain activity? A review on the learning of regulation of brain activity and a contribution to the discussion on specificity." *Front. Hum. Neurosci.*, vol. 9, p. 135, mar 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25870550

[70] S. McCarthy-Jones, "Taking Back the Brain: Could Neurofeedback Training Be Effective for Relieving Distressing Auditory Verbal Hallucinations in Patients With Schizophrenia?" *Schizophr. Bull.*, vol. 38, no. 4, pp. 678–682, jul 2012. [Online]. Available: https://academic.oup.com/schizophreniabulletin/article-lookup/doi/10.1093/schbul/sbs006

[71] S. Ruiz, N. Birbaumer, and R. Sitaram, "Abnormal Neural Connectivity in Schizophrenia and fMRI-Brain-Computer Interface as a Potential Therapeutic Approach," *Front. Psychiatry*, vol. 4, no. MAR, pp. 1–14, 2013. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fpsyt.2013.00017/abstract

[72] J. S. Cordes, K. A. K. Mathiak, M. Dyck, E. M. Alawi, T. J. Gaber, F. D. Zepf, M. Klasen, M. Zvyagintsev, R. C. Gur, and K. A. K. Mathiak, "Cognitive and neural strategies during control of the anterior cingulate cortex by fMRI neurofeedback in patients with schizophrenia." *Front. Behav. Neurosci.*, vol. 9, p. 169, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26161073

[73] M. Grosse-Wentrup, D. Mattia, and K. Oweiss, "Using brain-computer interfaces to induce neural plasticity and restore function." *J. Neural Eng.*, vol. 8, no. 2, p. 025004, 2011. [Online]. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4515347{&}tool=pmcentrez{&}rendertype=abstract

[74] N. Mrachacz-Kersting, N. Jiang, A. J. T. Stevenson, I. K. Niazi, V. Kostic, A. Pavlovic, S. Radovanovic, M. Djuric-Jovicic, F. Agosta, K. Dremstrup, and D. Farina, "Efficient neuroplasticity induction in chronic stroke patients by an associative brain-computer interface," *J. Neurophysiol.*, vol. 115, no. 3, pp. 1410–1421, mar 2016. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26719088http://www.physiology.org/doi/10.1152/jn.00918.2015

[75] K. K. Ang, K. S. G. Chua, K. S. Phua, C. Wang, Z. Y. Chin, C. W. K. Kuah, W. Low, and C. Guan, "A Randomized Controlled Trial of EEG-Based Motor Imagery Brain-Computer Interface Robotic Rehabilitation for Stroke." *Clin. EEG Neurosci.*, vol. 46, no. 4, pp. 310–20, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24756025

[76] A. Caria, C. Weber, D. Brötz, A. Ramos, L. F. Ticini, A. Gharabaghi, C. Braun, and N. Birbaumer, "Chronic stroke recovery after combined BCI training and physiotherapy: A case report," *Psychophysiology*, vol. 48, no. 4, pp. 578–582, 2011. [Online]. Available: http://doi.wiley.com/10.1111/j.1469-8986.2010.01117.x

[77] E. Mikołajewska and D. Mikołajewski, "The prospects of brain — computer interface applications in children," *Cent. Eur. J. Med.*, vol. 9, no. 1, pp. 74–79, 2013. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84897116587{&}partnerID=tZOtx3y1

[78] CDC, "Data & Statistics for Cerebral Palsy," 2016. [Online]. Available: http://www.cdc.gov/ncbddd/cp/data.html

[79] National Stroke Association, "Pediatric Stroke," 2016. [Online]. Available: http://www.stroke.org/understand-stroke/impact-stroke/pediatric-stroke

[80] CDC, "Rates of TBI-related Emergency Department Visits by Age Group - United States, 2001-2010," 2016. [Online]. Available: http://www.cdc.gov/traumaticbraininjury/data/rates{_}ed{_}byage.html

[81] F. Pichiorri, F. De Vico Fallani, F. Cincotti, F. Babiloni, M. Molinari, S. C. Kleih, C. Neuper, a. Kübler, and D. Mattia, "Sensorimotor rhythm-based brain–computer interface training: the impact on motor cortical responsiveness," *J. Neural Eng.*, vol. 8, no. 2, p. 025020, 2011. [Online]. Available: http://stacks.iop.org/1741-2552/8/i=2/a=025020?key=crossref.71f198e61c10309cd825e8bc321a39fa

[82] B. M. Young, Z. Nigogosyan, L. M. Walton, J. Song, V. A. Nair, S. W. Grogan, M. E. Tyler, D. F. Edwards, K. Caldera, J. A. Sattin, J. C. Williams, and V. Prabhakaran, "Changes in functional brain organization and behavioral correlations after rehabilitative therapy using a brain-computer interface," *Front. Neuroeng.*, vol. 7, p. 26, jul 2014. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fneng.2014.00026/abstract

[83] M. V. Johnston, "Clinical disorders of brain plasticity," *Brain Dev.*, vol. 26, no. 2, pp. 73–80, 2004.

[84] D. Boatman, J. Freeman, E. Vining, M. Pulsifer, D. Miglioretti, R. Minahan, B. Carson, J. Brandt, and G. McKhann, "Language recovery after left hemispherectomy in children with late-onset seizures," *Ann. Neurol.*, vol. 46, no. 4, pp. 579–586, 1999.

[85] S. V. Hiremath, W. Chen, W. Wang, S. Foldes, Y. Yang, E. C. Tyler-Kabara, J. L. Collinger, and M. L. Boninger, "Brain computer interface learning for systems based on electrocorticography and intracortical microelectrode arrays," *Front. Integr. Neurosci.*, vol. 9, jun 2015. [Online]. Available: http://journal.frontiersin.org/Article/10.3389/fnint.2015.00040/abstract

[86] A.-M. Brouwer, T. O. Zander, J. B. F. van Erp, J. E. Korteling, and A. W. Bronkhorst, "Using neurophysiological signals that reflect cognitive or affective state: six recommendations to avoid common pitfalls," *Front. Neurosci.*, vol. 9, Apr 2015. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnins.2015.00136/abstract

[87] E. Baque, L. Sakzewski, L. Barber, and R. N. Boyd, "Systematic review of phys iotherapy interventions to improve gross motor capacity and performance in children and adolescents with an acquired brain injury," *Brain Inj.*, vol. 9052, no. April, pp. 1–12, 2016. [Online]. Available: http://www.tandfonline.com/doi/full/10.3109/02699052.2016.1147079

[88] F. Cincotti, D. Mattia, F. Aloise, S. Bufalari, G. Schalk, G. Oriolo, A. Cherubini, M. G. Marciani, and F. Babiloni, "Non invasive brain-computer interface system: towards its application as assistive technology," *Brain Res. Bull.*, vol. 75, no. 6, pp. 796–803, 2008.

[89] B. M. Young, Z. Nigogosyan, L. M. Walton, A. Remsik, J. Song, V. A. Nair, M. E. Tyler, D. F. Edwards, K. Caldera, J. A. Sattin, J. C. Williams, and V. Prabhakaran, "Dose-response relationships using brain–computer interface technology impact stroke rehabilitation," *Front. Hum. Neurosci.*, vol. 9, jun 2015. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnhum.2015.00361/abstracthttp://journal.frontiersin.org/Article/10.3389/fnhum.2015.00361/abstract

[90] L. Snider, A. Majnemer, and V. Darsaklis, "Virtual reality as a therapeutic modality for children with cerebral palsy." *Dev. Neurorehabil.*, vol. 13, no. 2, pp. 120–128, 2010.

[91] P. R. Penn, F. D. Rose, and D. a. Johnson, "Virtual enriched environments in paediatric neuropsychological rehabilitation following traumatic brain injury: Feasibility, benefits and challenges." *Dev. Neurorehabil.*, vol. 12, no. 1, pp. 32–43, 2009.

[92] "Mindflex Duel." [Online]. Available: http://store.neurosky.com/products/mindflex-duel

[93] "Puzzlebox Orbit: Brain-Controlled Helicopter — Puzzlebox." [Online]. Available: https://puzzlebox.io/orbit/

[94] "Neurocog - : Neurocog." [Online]. Available: http://neurocog.com.au/

[95] "Star Wars Science - Force Trainer: Toys & Games Amazon.com." [Online]. Available: https://www.amazon.com/Star-Wars-Science-Force-Trainer/dp/B001UZHASY

[96] D. Kraus, G. Naros, R. Bauer, M. T. Leão, U. Ziemann, and A. Gharabaghi, "Brain-robot interface driven plasticity: Distributed modulation of corticospinal excitability." *Neuroimage*, vol. 125, pp. 522–532, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811915009799

[97] J. Faller, R. Scherer, U. Costa, E. Opisso, J. Medina, and G. R. Müller-Putz, "A Co-Adaptive Brain-Computer Interface for End Users with Severe Motor Impairment," *PLoS One*, vol. 9, no. 7, p. e101168, 2014. [Online]. Available: http://dx.plos.org/10.1371/journal.pone.0101168

[98] I. Daly, M. Billinger, J. Laparra-Hernández, F. Aloise, M. L. García, J. Faller, R. Scherer, and G. Müller-Putz, "On the control of brain-computer interfaces by users with cerebral palsy," *Clin. Neurophysiol.*, vol. 124, no. 9, pp. 1787–1797, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.clinph.2013.02.118

[99] D. J. McFarland, W. a. Sarnacki, T. M. Vaughan, and J. R. Wolpaw, "Brain-computer interface (BCI) operation: signal and noise during early training sessions," *Clin. Neurophysiol.*, vol. 116, no. 1, pp. 56–62, 2005. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245704002664

[100] B. Allison and C. Neuper, *Brain-Computer Interfaces*, ser. Human-Computer Interaction Series, D. S. Tan and A. Nijholt, Eds. London: Springer London, 2010. [Online]. Available: http://link.springer.com/10.1007/978-3-642-02091-9

[101] B. Allison, T. Lüth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Gräser, "BCI demographics: How many (and what kinds of) people can use an SSVEP BCI?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 107–116, 2010.

[102] I. Volosyak, D. Valbuena, T. Lüth, T. Malechka, and A. Gräser, "BCI demographics II: How many (and What Kinds of) people can use a high-frequency SSVEP BCI?" *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 3, pp. 232–239, 2011.

[103] J. C. Powers, K. Bieliaieva, S. Wu, C. S. Nam, and S. D. Meriney, "The Human Factors and Ergonomics of P300-Based Brain-Computer Interfaces," *Brain Sci. Brain Sci*, vol. 5, no. 5, pp. 318–354, 2015. [Online]. Available: https://www.mdpi.com/2076-3425/5/3/318

[104] I. Daly, J. Faller, R. Scherer, C. M. Sweeney-Reed, S. J. Nasuto, M. Billinger, and G. R. Müller-Putz, "Exploration of the neural correlates of cerebral palsy for

sensorimotor BCI control," *Front. Neuroeng.*, vol. 7, jul 2014. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fneng.2014.00020/abstract

[105] Y. C. Hung and A. M. Gordon, "Motor learning of a bimanual task in children with unilateral cerebral palsy," *Res. Dev. Disabil.*, vol. 34, no. 6, pp. 1891–1896, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.ridd.2013.03.008

[106] C. Crajé, M. van Elk, M. Beeren, H. T. van Schie, H. Bekkering, and B. Steenbergen, "Compromised motor planning and Motor Imagery in right Hemiparetic Cerebral Palsy," *Res. Dev. Disabil.*, vol. 31, no. 6, pp. 1313–1322, 2010.

[107] M. van Elk, C. Crajé, M. E. G. V. Beeren, B. Steenbergen, H. T. van Schie, and H. Bekkering, "Neural Evidence for Compromised Motor Imagery in Right Hemiparetic Cerebral Palsy," *Front. Neurol.*, vol. 1, no. 6, pp. 1313–1322, 2010. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fneur.2010.00150/abstract

[108] D. Plass-Oude Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, and D. Heylen, "Brain-Computer Interfacing and Games," in *Brain Comput. Interfaces*, 1st ed., ser. The Frontiers Collection, B. Graimann, G. Pfurtscheller, and B. Allison, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, ch. 10, pp. 149–178. [Online]. Available: http://link.springer.com/10.1007/978-1-84996-272-8{_}10

[109] K. Y. Manning, D. Fehlings, R. Mesterman, J. W. Gorter, L. Switzer, C. Campbell, and R. S. Menon, "Resting State and Diffusion Neuroimaging Predictors of Clinical Improvements Following Constraint-Induced Movement Therapy in Children With Hemiplegic Cerebral Palsy." *J. Child Neurol.*, vol. 30, no. 11, pp. 1507–14, 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25762587

[110] W. Kułak, W. Sobaniec, J.-S. Kuzia, and L. Boćkowski, "Neurophysiologic and neuroimaging studies of brain plasticity in children with spastic cerebral palsy." *Exp. Neurol.*, vol. 198, no. 1, pp. 4–11, 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16414042

[111] D. Tan and A. Nijholt, "Brain-Computer Interfaces and Human-Computer Interaction," in *Brain Comput. Interfaces*, 1st ed., B. Graimann, G. Pfurtscheller, and B. Allison, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, ch. 1, pp. 3–19. [Online]. Available: http://link.springer.com/10.1007/978-3-642-02091-9

[112] S. Moghimi, A. Kushki, A. Marie Guerguerian, and T. Chau, "A Review of EEG-Based Brain-Computer Interfaces as Access Pathways for Individuals with Severe Disabilities," *Assist. Technol.*, vol. 25, no. 2, pp. 99–110, 2013. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10400435.2012.723298

[113] C. G. Lim, T. S. Lee, C. Guan, D. S. S. Fung, Y. Zhao, S. S. W. Teng, H. Zhang, and K. R. R. Krishnan, "A Brain-Computer Interface Based Attention Training Program for Treating Attention Deficit Hyperactivity Disorder," *PLoS One*, vol. 7, no. 10, p. e46692, 2012. [Online]. Available: http://dx.plos.org/10.1371/journal.pone.0046692

[114] D. A. Rohani, H. B. D. Sorensen, and S. Puthusserypady, "Brain-computer interface using P300 and virtual reality: A gaming approach for treating ADHD," in *2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2014. IEEE, aug 2014, pp. 3606–3609. [Online]. Available: http://ieeexplore.ieee.org/document/6944403/

[115] J. Ehlers, D. Valbuena, A. Stiller, and A. Gräser, "Age-specific mechanisms in an SSVEP-based BCI scenario: evidences from spontaneous rhythms and neuronal oscillators." *Comput. Intell. Neurosci.*, vol. 2012, p. 967305, 2012. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84871373244{&}partnerID=tZOtx3y1

[116] J. D. Breshears, C. M. Gaona, J. L. Roland, M. Sharma, N. R. Anderson, D. T. Bundy, Z. V. Freudenburg, M. D. Smyth, J. Zempel, D. D. Limbrick, W. D. Smart, and E. C. Leuthardt, "Decoding motor signals from the pediatric cortex: implications for brain-computer interfaces in children." *Pediatrics*, vol. 128, no. 1, pp. e160–8, jul 2011. [Online]. Available: http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2010-1519

[117] J. Roland, K. Miller, Z. Freudenburg, M. Sharma, M. Smyth, C. Gaona, J. Breshears, M. Corbetta, and E. C. Leuthardt, "The effect of age on human motor electrocorticographic signals and implications for brain–computer interface applications," *J. Neural Eng.*, vol. 8, no. 4, p. 046013, 2011. [Online]. Available: http://stacks.iop.org/1741-2552/8/i=4/a=046013?key=crossref.c6375285f47d70d930db1b51e00d2661

[118] M. Ahn, M. Lee, J. Choi, and S. C. Jun, "A Review of Brain-Computer Interface Games and an Opinion Survey from Researchers, Developers and Users," *Sensors*, vol. 14, no. 8, pp. 14 601–14 633, 2014. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25116904

[119] S. J. Segalowitz, D. L. Santesso, and M. K. Jetha, "Electrophysiological changes during adolescence: A review," *Brain Cogn.*, vol. 72, no. 1, pp. 86–100, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.bandc.2009.10.003

[120] A. Chiang, C. Rennie, P. Robinson, S. van Albada, and C. Kerr, "Age trends and sex differences of alpha rhythms including split alpha peaks," *Clin. Neurophysiol.*, vol. 122, no. 8, pp. 1505–1517, 2011. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245711000903

[121] Y. Liu, H. Zhang, M. Chen, and L. Zhang, "A Boosting-Based Spatial-Spectral Model for Stroke Patients' EEG Analysis in Rehabilitation Training." *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 169–79, jan 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7214285/

[122] H. Ji, J. Li, R. Lu, R. Gu, L. Cao, and X. Gong, "EEG Classification for Hybrid Brain-Computer Interface Using a Tensor Based Multiclass Multimodal Analysis Scheme," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–15, 2016. [Online]. Available: http://www.hindawi.com/journals/cin/2016/1732836/

[123] J. Li, L. Zhang, D. Tao, H. Sun, and Q. Zhao, "A Prior neurophysiologic knowledge free tensor-based scheme for single trial EEG classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 2, pp. 107–115, 2009.

[124] Y. Zhang, Q. Zhao, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Removal of EEG artifacts for BCI applications using fully Bayesian tensor completion," in *2016 IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, mar 2016, pp. 819–823. [Online]. Available: http://ieeexplore.ieee.org/document/7471789/

[125] Huijuan Yang, S. Sakhavi, Kai Keng Ang, and Cuntai Guan, "On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification," in *2015 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2015-Novem. IEEE, aug 2015, pp. 2620–2623. [Online]. Available: http://ieeexplore.ieee.org/document/7318929/

[126] E. Santana, A. J. Brockmeier, and J. C. Principe, "Joint optimization of algorithmic suites for EEG analysis," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2014, pp. 2997–3000, 2014.

[127] C. Grozea, C. D. Voinescu, and S. Fazli, "Bristle-sensors–low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications." *J. Neural Eng.*, vol. 8, no. 2, p. 025008, 2011.

[128] C. Guger, G. Krausz, B. Z. Allison, and G. Edlinger, "Comparison of dry and gel based electrodes for P300 brain-computer interfaces," *Front. Neurosci.*, vol. 6, no. MAY, pp. 1–7, 2012.

[129] T. M. Lau, "How Many Electrodes Are Really Needed for EEG-Based Mobile Brain Imaging?" *J. Behav. Brain Sci.*, vol. 02, no. 03, pp. 387–393, 2012.

[130] F. Nijboer, J. Clausen, B. Z. Allison, and P. Haselager, "The Asilomar Survey: Stakeholders' Opinions on Ethical Issues Related to Brain-Computer Interfacing," *Neuroethics*, vol. 6, no. 3, pp. 541–578, 2013. [Online]. Available: http://link.springer.com/10.1007/s12152-011-9132-6

[131] E. a. Felton, J. C. Williams, G. C. Vanderheiden, and R. G. Radwin, "Mental workload during brain–computer interface training," *Ergonomics*, vol. 55, no. 5, pp. 526–537, 2012.

[132] X. Qian, B. R. Y. Loo, F. X. Castellanos, S. Liu, H. L. Koh, X. W. W. Poh, R. Krishnan, D. Fung, M. W. Chee, C. Guan, T.-S. Lee, C. G. Lim, and J. Zhou, "Brain-computer-interface-based

intervention re-normalizes brain functional network topology in children with attention deficit/hyperactivity disorder," *Transl. Psychiatry*, vol. 8, no. 1, p. 149, dec 2018. [Online]. Available: http://www.nature.com/articles/s41398-018-0213-8

[133] E.-S. McLaren and A. N. Antle, "Exploring and Evaluating Sound for Helping Children Self-Regulate with a Brain-Computer Application," in *Proc. 2017 Conf. Interact. Des. Child. - IDC '17*, 2017, pp. 393–398. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3078072.3084299

[134] H. Gurkok, A. Nijholt, and M. Poel, "Brain-Computer Interface Games: Towards a Framework," in *Handb. Digit. Games Entertain. Technol.* Singapore: Springer Singapore, 2017, pp. 133–150. [Online]. Available: http://link.springer.com/10.1007/978-981-4560-50-4{_}5

[135] J. J. S. Norton, J. Mullins, B. E. Alitz, and T. Bretl, "The performance of 9–11-year-old children using an SSVEP-based BCI for target selection," *J. Neural Eng.*, vol. 15, no. 5, p. 056012, oct 2018. [Online]. Available: http://stacks.iop.org/1741-2552/15/i=5/a=056012?key=crossref.8f8d820e11d1c1ae3c436823195ec8f4

[136] J. M. De Oliveira, R. C. G. Fernandes, C. S. Pinto, P. R. Pinheiro, S. Ribeiro, and V. H. C. De Albuquerque, "Novel virtual environment for alternative treatment of children with cerebral palsy," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–10, jun 2016. [Online]. Available: https://www.hindawi.com/journals/cin/2016/8984379/

[137] "Seventh International BCI Meeting- BCIs: Not Getting Lost in Translation." [Online]. Available: http://bcisociety.org/wp-content/uploads/2018/05/BCI2018AbstractBook-final.pdf

[138] R. Mouček, L. Vařeka, T. Prokop, J. Štěbeták, and P. Brůha, "Event-related potential data from a guess the number brain-computer interface experiment on school children," *Sci. Data*, vol. 4, p. 160121, mar 2017. [Online]. Available: http://www.nature.com/articles/sdata2016121

[139] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res. Rev.*, vol. 29, no. 2-3, pp. 169–195, apr 1999. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0165017398000563

[140] H. M.B., S. R., V. K., Y. M., M. A., S. J., and C. R.F., "NEUROdevelopment in PReschool Children Of FIfe and Lothian Epilepsy Study: NEUROPROFILES - A population-based study," pp. 56–57, 2015. [Online]. Available: http://ovidsp.ovid.com/ovidweb.cgi?T=JS{&}PAGE=reference{&}D=emed12{&}NEWS=N{&}AN=71754114

[141] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, 2009. [Online]. Available: http://hdl.handle.net/1721.1/54669

[142] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, jun 2000. [Online]. Available: http://circ.ahajournals.org/cgi/doi/10.1161/01.CIR.101.23.e215

[143] R. Bro and U. V. Amsterdam, "Multi-way Analysis in the Food Industry," *Food Technol.*, no. november, 1998.

[144] E. Frank, M. A. Hall, and I. H. Witten, "The WEKA Workbench," in *Morgan Kaufmann, Fourth Ed.* Elsevier, 2016, pp. 553–571. [Online]. Available: https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf

[145] J. Hill, T. N. Lal, M. Schroeder, T. Hinterberger, G. Widman, C. Elger, B. Schölkopf, and N. Birbaumer, "Classifying Event-Related Desynchronization in EEG, ECoG and MEG signals," *Pattern Recognit.*, vol. 4174, pp. 404–413, 2006. [Online]. Available: https://link.springer.com/chapter/10.1007/11861898_41

[146] L. J. P. Van Der Maaten and G. E. Hinton, "Visualizing high-dimensional data using t-sne," *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008. [Online]. Available: http://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf

[147] S. Haufe and A. Ewald, "A Simulation Framework for Benchmarking EEG-Based Brain Connectivity Estimation Methodologies," *Brain Topogr.*, pp. 1–18, jun 2016. [Online]. Available: http://link.springer.com/10.1007/s10548-016-0498-y

[148] A. Birca, L. Carmant, A. Lortie, P. Vannasing, H. Sauerwein, M. Robert, L. Lemay, X.-P. Wang, D. Piper, and V. Donici, "Maturational changes of 5Hz SSVEPs elicited by intermittent photic stimulation," *Int. J. Psychophysiol.*, vol. 78, no. 3, pp. 295–298, dec 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167876010006926

[149] H. Wang, Y. Zhang, N. R. Waytowich, D. J. Krusienski, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Discriminative Feature Extraction via Multivariate Linear Regression for SSVEP-based BCI," *Neural Syst. Rehabil. Eng. IEEE Trans.*, vol. PP, no. 99, p. 1, 2016.

[150] R. Zink, B. Hunyadi, S. V. Huffel, and M. D. Vos, "Tensor-based classification of an auditory mobile BCI without a subject-specific calibration phase," *J. Neural Eng.*, vol. 13, no. 2, p. 026005, apr 2016. [Online]. Available: http://iopscience.iop.org/article/10.1088/1741-2560/13/2/026005/pdf

[151] R. G. O'Connell, P. M. Dockree, and S. P. Kelly, "A supramodal accumulation-to-bound signal that determines perceptual decisions in humans," *Nat. Neurosci.*, vol. 15, no. 12, pp. 1729–1735, dec 2012. [Online]. Available: http://www.nature.com/articles/nn.3248

[152] A. H. Phan, P. Tichavsky, and A. Cichocki, "CANDECOMP/PARAFAC decomposition of high-order tensors through tensor reshaping," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4847–4860, oct 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6542035/

[153] L. Sorber, M. Van Barel, and L. De Lathauwer, "Structured Data Fusion," *IEEE J. Sel. Top. Signal Process.*, vol. 9, no. 4, pp. 586–600, 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7039240/

[154] E. Acar, T. G. Kolda, and D. M. Dunlavy, "All-at-once Optimization for Coupled Matrix and Tensor Factorizations," *Mlg'11*, no. 1, may 2011. [Online]. Available: http://arxiv.org/abs/1105.3422

[155] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs, and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," *Chemom. Intell. Lab. Syst.*, vol. 129, pp. 53–63, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.chemolab.2013.06.006

[156] E. Acar, E. E. Papalexakis, G. Gürdeniz, M. a. Rasmussen, A. J. Lawaetz, M. Nilsson, and R. Bro, "Structure-revealing data fusion," *BMC Bioinformatics*, vol. 15, no. 1, p. 239, 2014. [Online]. Available: http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-15-239

[157] E. Acar, R. Bro, and A. K. Smilde, "Data Fusion in Metabolomics Using Coupled Matrix and Tensor Factorizations," *Proc. IEEE*, vol. 103, no. 9, pp. 1602–1620, 2015.

[158] G. Augusto Chiarenza, D. Papakostopoulos, F. Giordana, and A. Guareschi-Cazzullo, "Movement-related brain macropotentials during skilled performances. A developmental study," *Electroencephalogr. Clin. Neurophysiol.*, vol. 56, no. 4, pp. 373–383, 1983. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0013469483902638

[159] K. Cuevas, V. Raj, and M. A. Bell, "A frequency band analysis of two-year-olds' memory processes," *Int. J. Psychophysiol.*, vol. 83, no. 3, pp. 315–322, 2012.

[160] A. Shield, K. Knapke, M. Henry, S. Srinivasan, and A. Bhat, "Impaired praxis in gesture imitation by deaf children with autism spectrum disorder," *Autism Dev. Lang. Impair.*, vol. 2, p. 239694151774567, 2017. [Online]. Available: http://journals.sagepub.com/doi/pdf/10.1177/2396941517745674

[161] D. Wechsler, *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson, 2011. [Online]. Available: https://www.pearsonclinical.co.uk/Psychology/AdultCognitionNeuropsychologyandLanguage/AdultGeneralAbilities/WASI-II/WechslerAbbreviatedScaleofIntelligenceSecondEdition.aspx

[162] ——, *Wechsler Individual Achievement Test - Second Edition Abbreviated (WIAT-IIA)*, 2nd ed. San Antonio, TX: Psychological Corporation., 2005. [Online]. Available: http://www.worldcat.org/title/wiat-ii-abbreviated-wechsler-individual-achievement-test/oclc/57531880

[163] U. Ravens-Sieberer and M. Bullinger, "Assessing health-related quality of life in chronically ill children with the German KINDL: first psychometric and content analytical results," *Qual. Life Res.*, vol. 7, no. 5, pp. 399–407, jul 1998. [Online]. Available: http://link.springer.com/10.1023/A:1008853819715

[164] M. M. Pangelinan, F. A. Kagerer, B. Momen, B. D. Hatfield, and J. E. Clark, "Electrocortical dynamics reflect age-related differences in movement kinematics among children and adults," *Cereb. Cortex*, vol. 21, no. 4, pp. 737–747, apr 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/20805237

[165] S. Bender, M. Weisbrod, H. Bornfleth, F. Resch, and R. Oelkers-Ax, "How do children prepare to react? Imaging maturation of motor preparation and stimulus anticipation by late contingent negative variation," *Neuroimage*, vol. 27, no. 4, pp. 737–752, 2005. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16027009

[166] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," 2016. [Online]. Available: https://www.tensorlab.net/

[167] N. Vervliet, O. Debals, L. Sorber, and L. De Lathauwer, "Breaking the curse of dimensionality using decompositions of incomplete tensors: Tensor-based scientific computing in big data analysis," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 71–79, sep 2014. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6879619

[168] G. Tomasi and R. Bro, "A comparison of algorithms for fitting the PARAFAC model," *Comput. Stat. Data Anal.*, vol. 50, no. 7, pp. 1700–1734, 2006.

[169] N. Li, S. Kindermann, and C. Navasca, "Some convergence results on the Regularized Alternating Least-Squares method for tensor decomposition," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 796–812, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.laa.2011.12.002

[170] G. Zhou, A. Cichocki, Y. Zhang, and D. P. Mandic, "Group Component Analysis for Multiblock Data: Common and Individual Feature Extraction," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 11, pp. 2426–2439, 2016. [Online]. Available: http://arxiv.org/abs/1212.3913

[171] M. Berchicci, T. Zhang, L. Romero, a. Peters, R. Annett, U. Teuscher, M. Bertollo, Y. Okada, J. Stephen, and S. Comani, "Development of Mu Rhythm in Infants and Preschool Children," *Dev. Neurosci.*, vol. 33, no. 2, pp. 130–143, 2011. [Online]. Available: http://www.karger.com/doi/10.1159/000329095

[172] E. Kinney-Lang, M. Yoong, M. Hunter, K. K. Tallur, J. Shetty, A. McLellan, R. F. Chin, and J. Escudero, "Analysis of EEG networks and their correlation with cognitive impairment in preschool children with epilepsy," *bioRxiv*, p. 237172, sep 2018. [Online]. Available: https://www.biorxiv.org/content/early/2018/09/28/237172

[173] B. S. Chang and D. H. Lowenstein, "Epilepsy," *N. Engl. J. Med.*, vol. 349, no. 13, pp. 1257–1266, sep 2003. [Online]. Available: http://www.nejm.org/doi/abs/10.1056/NEJMra022308

[174] C. Reilly, P. Atkinson, K. B. Das, R. F. M. C. Chin, S. E. Aylett, V. Burch, C. Gillberg, R. C. Scott, and B. G. R. Neville, "Neurobehavioral Comorbidities in Children With Active Epilepsy: A Population-Based Study," *Pediatrics*, vol. 133, no. 6, pp. e1586–e1593, jun 2014. [Online]. Available: http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2013-3787

[175] C. Reilly, P. Atkinson, K. B. Das, R. F. Chin, S. E. Aylett, V. Burch, C. Gillberg, R. C. Scott, and B. G. Neville, "Factors associated with quality of life in active childhood epilepsy: A population-based study," *Eur. J. Paediatr. Neurol.*, vol. 19, no. 3, pp. 308–313, may 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1090379815000069

[176] E.-H. Kim and T.-S. Ko, "Cognitive impairment in childhood onset epilepsy: up-to-date information about its causes," *Korean J. Pediatr.*, vol. 59, no. 4, p. 155, apr 2016. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/27186225

[177] K. Rantanen, K. Eriksson, and P. Nieminen, "Cognitive impairment in preschool children with epilepsy," *Epilepsia*, vol. 52, no. 8, pp. 1499–1505, aug 2011. [Online]. Available: http://doi.wiley.com/10.1111/j.1528-1167.2011.03092.x

[178] M. Yoong, M. Hunter, J. Stephen, A. Quigley, J. Jones, J. Shetty, A. McLellan, M. E. Bastin, and R. F. Chin, "Cognitive impairment in early onset epilepsy is associated with reduced left thalamic volume," *Epilepsy Behav.*, vol. 80, pp. 266–271, mar 2018. [Online]. Available: https://www.epilepsybehavior.com/article/S1525-5050(17)30945-9/pdf

[179] D. B. Bailey, "Critical thinking about critical periods," 2001.

[180] W. A. Hauser, J. F. Annegers, and W. A. Rocca, "Descriptive Epidemiology of Epilepsy: Contributions of Population-Based Studies From Rochester, Minnesota," *Mayo Clin. Proc.*, vol. 71, no. 6, pp. 576–586, jun 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0025619611641153

[181] B. Neville, "Epilepsy in childhood." *BMJ Br. Med. J.*, vol. 315, pp. 924–930, 1997. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2127609/pdf/9361544.pdf

[182] M. J. England, C. T. Liverman, A. M. Schultz, and L. M. Strawbridge, "Epilepsy across the spectrum: Promoting health and understanding. A summary of the Institute of Medicine report," pp. 266–276, 2012.

[183] A. R. Brooks-Kayal, K. G. Bath, A. T. Berg, A. S. Galanopoulou, G. L. Holmes, F. E. Jensen, A. M. Kanner, T. J. O'Brien, V. H. Whittemore, M. R. Winawer, M. Patel, and H. E. Scharfman, "Issues related to symptomatic and disease-modifying treatments affecting cognitive and neuropsychiatric comorbidities of epilepsy," *Epilepsia*, vol. 54, no. SUPPL.4, pp. 44–60, aug 2013. [Online]. Available: http://doi.wiley.com/10.1111/epi.12298

[184] M. Yoong, "Quantifying the deficit-imaging neurobehavioural impairment in childhood epilepsy." *Quant. Imaging Med. Surg.*, vol. 5, no. 2, pp. 225–37, apr 2015. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/25853081

[185] M. T. Horstmann, S. Bialonski, N. Noennig, H. Mai, J. Prusseit, J. Wellmer, H. Hinrichs, and K. Lehnertz, "State dependent properties of epileptic brain networks: Comparative graph-theoretical analyses of simultaneously recorded EEG and MEG," *Clin. Neurophysiol.*, vol. 121, no. 2, pp. 172–185, feb 2010. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245709006154

[186] E. van Diessen, W. M. Otte, C. J. Stam, K. P. Braun, and F. E. Jansen, "Electroencephalography based functional networks in newly diagnosed childhood epilepsies," *Clin. Neurophysiol.*, vol. 127, no. 6, pp. 2325–2332, 2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245716001085

[187] E. Bullmore and O. Sporns, "Erratum: Complex brain networks: graph theoretical analysis of structural and functional systems," *Nat. Rev. Neurosci.*, vol. 10, no. 4, pp. 312–312, apr 2009. [Online]. Available: http://www.nature.com/articles/nrn2618

[188] M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," *J. Am. Stat. Assoc.*, vol. 32, no. 200, p. 675, dec 1937. [Online]. Available: https://www.jstor.org/stable/2279372?origin=crossref

[189] P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch, "EEG coherency," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 5, pp. 499–515, nov 1997. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0013469497000667

[190] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, "Identifying true brain interaction from EEG data using the imaginary part of coherency," *Clin. Neurophysiol.*, vol. 115, no. 10, pp. 2292–2307, 2004.

[191] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, and K.-R. Müller, "Robustly Estimating the Flow Direction of Information in Complex Physical

Systems," *Phys. Rev. Lett.*, vol. 100, no. 23, p. 234101, jun 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.100.234101

[192] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Hum. Brain Mapp.*, vol. 28, no. 11, pp. 1178–1193, 2007.

[193] M. Vinck, R. Oostenveld, M. Van Wingerden, F. Battaglia, and C. M. A. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *Neuroimage*, vol. 55, no. 4, pp. 1548–1565, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2011.01.055

[194] S. Haufe, V. V. Nikulin, K.-R. Müller, and G. Nolte, "A critical assessment of connectivity measures for EEG data: A simulation study," *Neuroimage*, vol. 64, no. 1, pp. 120–133, jan 2013. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1053811912009469

[195] P. Tewarie, E. van Dellen, A. Hillebrand, and C. J. Stam, "The minimum spanning tree: An unbiased method for brain network analysis," *Neuroimage*, vol. 104, pp. 177–188, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2014.10.015

[196] Fornito, "Network scaling effects in graph analytic studies of human resting-state fMRI data," *Front. Syst. Neurosci.*, vol. 4, p. 22, jun 2010. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnsys.2010.00022/abstract

[197] B. C. M. van Wijk, C. J. Stam, and A. Daffertshofer, "Comparing brain networks of different size and connectivity density using graph theory," *PLoS One*, vol. 5, no. 10, p. e13701, 2010. [Online]. Available: http://www.nwo.nlhttp://dx.plos.org/10.1371/journal.pone.0013701

[198] K. Smith, H. Azami, M. A. Parra, J. M. Starr, and J. Escudero, "Cluster-span threshold: An unbiased threshold for binarising weighted complete networks in functional connectivity analysis," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2015-Novem, pp. 2840–2843, 2015.

[199] J. B. Kruskal, "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem," *Proc. Am. Math. Soc.*, vol. 7, no. 1, p. 48, feb 1956. [Online]. Available: https://www.jstor.org/stable/2033241?origin=crossref

[200] K. Smith, D. Abasolo, and J. Escudero, "Accounting for the complex hierarchical topology of EEG phase-based functional connectivity in network binarisation," *PLoS One*, 2017.

[201] A. R. Gilpin, "Table for Conversion of Kendall's Tau to Spearman's Rho Within the Context of Measures of Magnitude of Effect for Meta-Analysis," *Educ. Psychol. Meas.*, vol. 53, no. 1, pp. 87–92, mar 1993. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0013164493053001007

[202] N. Shong, "Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data," Ph.D. dissertation, University of Pittsburgh, 2010. [Online]. Available: http://d-scholarship.pitt.edu/8056/

[203] G. Fraga González, M. Van der Molen, G. Žarić, M. Bonte, J. Tijms, L. Blomert, C. Stam, and M. Van der Molen, "Graph analysis of EEG resting state functional networks in dyslexic readers," *Clin. Neurophysiol.*, vol. 127, no. 9, pp. 3165–3175, 2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245716304539

[204] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *2014 Sci. Inf. Conf.* IEEE, aug 2014, pp. 372–378. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6918213

[205] Z.-H. Zhou and X.-Y. Liu, "On multi-class cost-sensitive learning," *Comput. Intell.*, vol. 26, no. 3, pp. 232–257, 2010. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8640.2010.00358.x

[206] J. Shao, "Bootstrap Model Selection," *J. Am. Stat. Assoc.*, vol. 91, no. 434, pp. 655–665, 1996. [Online]. Available: https://www.jstor.org/stable/2291661?origin=crossref

[207] E. OREKHOVA, T. STROGANOVA, I. POSIKERA, and M. ELAM, "EEG theta rhythm in infants and preschool children," *Clin. Neurophysiol.*, vol. 117, no. 5, pp. 1047–1062, may 2006. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S1388245706000095

[208] T. a. B. Snijders, "The degree variance: An index of graph heterogeneity," *Soc. Networks*, vol. 3, no. 3, pp. 163–174, jan 1981. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0378873381900149

[209] A. Amador, P. Valdés Sosa, R. Pascual Marqui, L. Garcia, R. Lirio, and J. Bayard, "On the structure of EEG development," *Electroencephalogr. Clin. Neurophysiol.*, vol. 73, no. 1, pp. 10–19, jul 1989. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/0013469489900151

[210] B. Liu, L. He, S. Zhe, Y. Li, and Z. Xu, "DeepCP : Flexible Nonlinear Tensor Decomposition," 2017. [Online]. Available: http://bayesiandeeplearning.org/2017/papers/26.pdf

[211] R. Schutski, J. Zhao, T. M. Henderson, and G. E. Scuseria, "Tensor-structured coupled cluster theory," *J. Chem. Phys.*, vol. 147, no. 18, p. 184113, nov 2017. [Online]. Available: https://arxiv.org/pdf/1708.02674.pdf

[212] A. Cichocki, A.-H. Phan, Q. Zhao, N. Lee, I. V. Oseledets, M. Sugiyama, and D. Mandic, "Tensor Networks for Dimensionality Reduction and Large-Scale Optimizations. Part 2 Applications and Future Perspectives," *Found. Trends® Mach. Learn.*, vol. 9, no. 6, pp. 249–429, 2017. [Online]. Available: http://www.nowpublishers.com/article/Details/MAL-067

[213] A. H. Phan, P. Tichavsky, and A. Cichocki, "Partitioned Hierarchical alternating least squares algorithm for CP tensor decomposition," in *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* IEEE, mar 2017, pp. 2542–2546. [Online]. Available: http://ieeexplore.ieee.org/document/7952615/

[214] I. Kisil, G. G. Calvi, and D. P. Mandic, "Tensor Valued Common and Individual Feature Extraction: Multi-dimensional Perspective," *arXiv*, 2017. [Online]. Available: http://arxiv.org/abs/1711.00487

[215] G. Zhou, Q. Zhao, Y. Zhang, T. Adali, S. Xie, and A. Cichocki, "Linked Component Analysis from Matrices to High-Order Tensors: Applications to Biomedical Data," *Proc. IEEE*, vol. 104, no. 2, pp. 310–331, 2016.

[216] Y. Liu, Q. Zhao, and L. Zhang, "Uncorrelated multiway discriminant analysis for motor imagery EEG classification." *Int. J. Neural Syst.*, vol. 25, no. 4, p. 1550013, 2015.

[217] G. Favier and A. L. de Almeida, "Overview of constrained PARAFAC models," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, p. 142, 2014. [Online]. Available: http://arxiv.org/abs/1405.7442http://asp.eurasipjournals.springeropen.com/articles/10.1186/1687-6180-2014-142

[218] B. Hunyadi, P. Dupont, W. Van Paesschen, and S. Van Huffel, "Tensor decompositions and data fusion in epileptic electroencephalography and functional magnetic resonance imaging data," p. e1197, jan 2017. [Online]. Available: http://doi.wiley.com/10.1002/widm.1197

[219] G. Deshpande, D. Rangaprakash, L. Oeding, A. Cichocki, and X. P. Hu, "A new generation of brain-computer interfaces driven by discovery of latent EEG-fMRI linkages using tensor decomposition," *Front. Neurosci.*, vol. 11, no. JUN, p. 246, jun 2017. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnins.2017.00246/full

[220] B. Cao, L. He, X. Wei, M. Xing, P. S. Yu, H. Klumpp, and A. D. Leow, "t-BNE: Tensor-based Brain Network Embedding," in *Proc. 2017 SIAM Int. Conf. Data Min.*, 2017, pp. 189–197. [Online]. Available: http://www.siam.org/journals/ojsa.phphttps://epubs.siam.org/doi/10.1137/1.9781611974973.22

[221] K. Naskovska, A. A. Korobkov, M. Haardt, and J. Haueisen, "Analysis of the Photic Driving Effect via joint EEG and MEG data processing based on the Coupled CP decomposition," in *Eusipco'17 Proc. 25th Eur. Signal Process. Conf.* IEEE, aug 2017, pp. 1325–1329. [Online]. Available: http://ieeexplore.ieee.org/document/8081415/

[222] S. Ferdowsi, V. Abolghasemi, and S. Sanei, "A new informed tensor factorization approach to EEG-fMRI fusion." *J. Neurosci. Methods*, vol. 254, pp. 27–35, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165027015002733

[223] F. Cong, A. H. Phan, H. Lyytinen, T. Ristaniemi, and A. Cichocki, "Classifying healthy children and children with attention deficit through features derived from sparse and nonnegative tensor factorization using event-related potential," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6365 LNCS, pp. 620–628, 2010.

[224] Shuo Zhou, "Accelerating Online CP Decompositions for Higher Order Tensors," *Kdd 2016*, p. 4503, 2016.

[225] E. M. Stoudenmire and D. J. Schwab, "Supervised Learning with Tensor Networks," in *30th Conf. Neural Inf. Process. Syst. (NIPS 2016)*, no. Nips, 2016. [Online]. Available: https://github.com/emstoudenmire/TNML.

[226] F. Wu, X. Tan, Y. Yang, D. Tao, S. Tang, and Y. Zhuang, "Supervised Nonnegative Tensor Factorization with Maximum-Margin Constraint," in *Proc. 27th AAAI Conf. Artif. Intell.*, 2013, pp. 962–968. [Online]. Available: www.aaai.org

[227] L. Spyrou and J. Escudero, "Graph regularised tensor factorisation of EEG signals based on network connectivity measures," in *2017 IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, mar 2017, pp. 944–948. [Online]. Available: http://ieeexplore.ieee.org/document/7952295/